



Know All About

Digital and Audio Books

Barney Everhart

Frank Rowan

First Edition, 2012

ISBN 978-81-323-1490-5

WWT

© All rights reserved.

Published by:

College Publishing House
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

WORLD TECHNOLOGIES

Table of Contents

Chapter 1 - Electronic Book

Chapter 2 - Comparison of e-book formats

Chapter 3 - Iliad

Chapter 4 - Amazon Kindle

Chapter 5 - PocketBook Reader

Chapter 6 - Introduction to Audiobook

Chapter 7 - Windows Media Audio

Chapter 8 - Advanced Audio Coding

Chapter 9 - DAISY Digital Talking Book

Chapter 10 - Audio Codecs

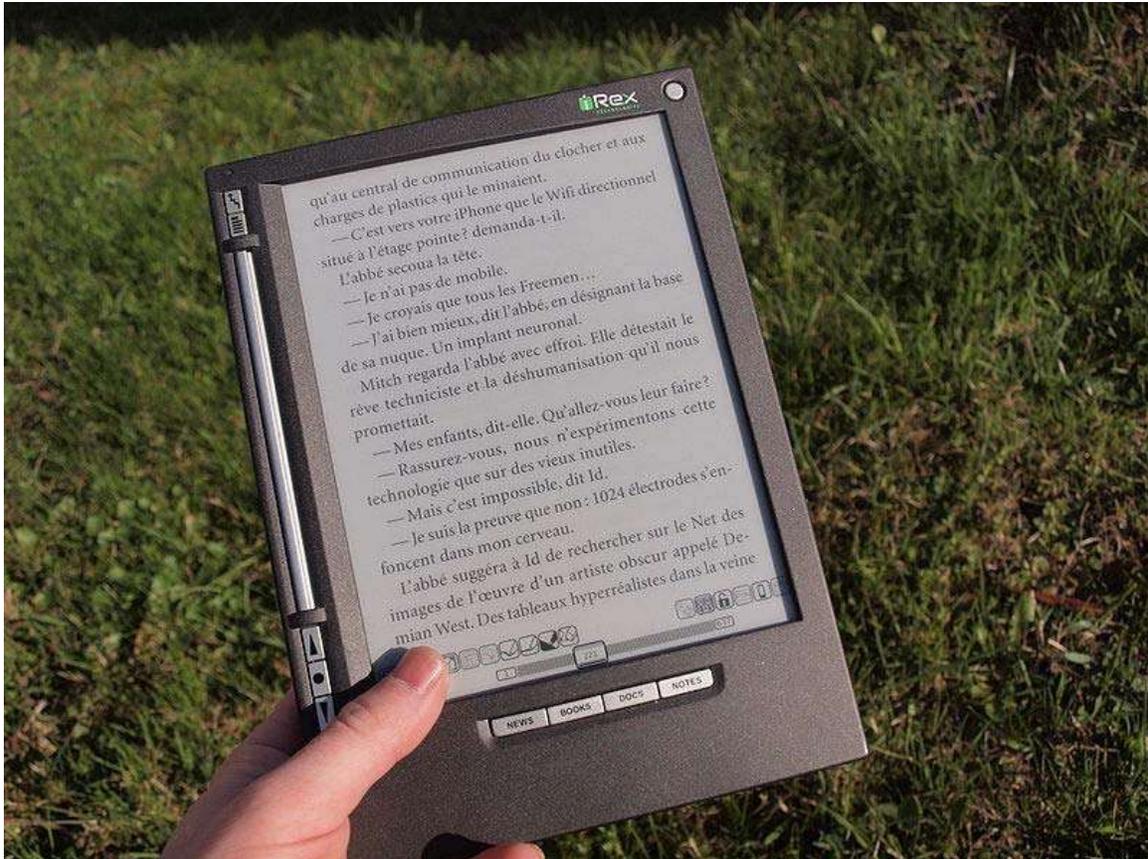
Chapter 1

Electronic Book



The books pictured above would use less than 1% of the device's memory

An **electronic book** (also **e-book**, **ebook**, **digital book**) is a text and image-based publication in digital form produced on, published by, and readable on computers or other digital devices. Sometimes the equivalent of a conventional printed book, e-books can also be born digital. The *Oxford Dictionary of English* defines the e-book as "an electronic version of a printed book," but e-books can and do exist without any printed equivalent. E-books are usually read on dedicated hardware devices known as *e-Readers* or *e-book devices*. Personal computers and some cell phones can also be used to read e-books.



iLiad in sunlight

History

Among the earliest general e-books were those in *Project Gutenberg*, started by Michael S. Hart in 1971. An early e-book implementation were the desktop prototypes for a proposed notebook computer, the *Dynabook*, in the 1970s at PARC, which would be a general-purpose portable personal computer, including reading books. Similar ideas were expressed at the same time by Paul Drucker.

Early e-books were generally written for specialty areas and a limited audience, meant to be read only by small and devoted interest groups. The scope of the subject matter of

these e-books included technical manuals for hardware, manufacturing techniques and other subjects. In the 1990s, the general availability of the Internet made transferring electronic files much easier, including e-books.



A user viewing an electronic page on an eBook-reading device



iLiad e-book reader equipped with e-paper display

Numerous e-book formats emerged and proliferated, some supported by major software companies such as Adobe with its PDF format, and others supported by independent and open-source programmers. Multiple readers followed multiple formats, most of them specializing in only one format, and thereby fragmenting the e-book market even more. Due to exclusiveness and limited readerships of e-books, the fractured market of independents and specialty authors lacked consensus regarding a standard for packaging and selling e-books. E-books continued to gain in their own underground markets. Many e-book publishers began distributing books that were in the public domain. At the same time, authors with books that were not accepted by publishers offered their works online so they could be seen by others. Unofficial (and occasionally unauthorized) catalogs of books became available over the web, and sites devoted to e-books began disseminating information about e-books to the public.

U.S. Libraries began providing free e-books to the public in 1998 through their web sites and associated services, although the e-books were primarily scholarly, technical or professional in nature, and could not be downloaded. In 2003, libraries began offering free downloadable popular fiction and non-fiction e-books to the public, launching an e-book lending model that worked much more successfully for public libraries. The number of library e-book distributors and lending models continued to increase over the next few years. In 2010, a Public Library Funding and Technology Access Study found that 66%

of public libraries in the U.S. were offering e-books, and a large movement in the library industry began seriously examining the issues related to lending e-books, acknowledging a tipping point of broad e-book usage.



As of 2009, new marketing models for e-books were being developed and dedicated reading hardware was produced. E-books (as opposed to ebook readers) have yet to achieve global distribution. In the United States, as of September 2009, the Amazon Kindle model and Sony's PRS-500 were the dominant ereading devices . By March 2010, some reported that the Barnes & Noble Nook may be selling more units than the Kindle. On January 27, 2010 Apple, Inc. launched a multi-function device called the iPad and announced agreements with five of the six largest publishers that would allow Apple to distribute e-books. However, many publishers and authors have not endorsed the concept of electronic publishing, citing issues with demand, piracy and proprietary devices.

In July 2010, online bookseller Amazon.com reported sales of ebooks for its proprietary Kindle outnumbered sales of hardcover books for the first time ever during the second quarter of 2010, saying it sold 140 e-books for every 100 hardcover books, including hardcovers for which there was no digital edition. In July this number had increased to 180 Kindle ebooks per 100 hardcovers . Paperback book sales are still much larger than either hardcover or e-book; the American Publishing Association estimated e-books represented 8.5% of sales as of mid-2010.

Timeline

- 1971: Michael S. Hart launches *Project Gutenberg*.

- 1985-1992 Robert Stein starts Voyager Company Expanded Books and books on CD-ROMs.
- 1992: Charles Stack's Book Stacks Unlimited begins selling new physical books online.
- 1993: Zahur Klemath Zapata develops the first software to read digital books. Digital book version 1 and the first digital book is published *On Murder Considered as one of the Fine Arts* (Thomas de Quincey).
- 1993: Digital Book, Inc. offers the first 50 digital books in Floppy disk with Digital Book Format (DBF).
- 1993: Hugo Award for Best Novel nominee texts published on CD-ROM by Brad Templeton.
- 1993: Bibliobytes, a project of free digital books online in Internet.
- 1995: Amazon starts to sell physical books in Internet.
- 1995: Online poet Alexis Kirke discusses the need for wireless internet electronic paper readers in his article "The Emuse".
- 1996: Project Gutenberg reaches 1,000 titles. The target is 1,000,000
- 1998 Kim Blagg obtained the first ISBN issued to an ebook and began marketing multimedia-enhanced ebooks on CDs through retailers including amazon.com, bn.com and borders.com. Shortly thereafter through her company "Books OnScreen" she introduced the ebooks at the Book Expo America in Chicago, IL to an impressed, but unconvinced bookseller audience.



- 1998: Launched the first ebook Readers: Rocket ebook and SoftBook.
- 1998: Cybook / Cybook Gen1 Sold and manufactured at first by Cytale (1998–2003) then by Bookeen
- 1998-1999: Websites selling ebooks in English, like *eReader.com* and *eReads.com*.
- 1999: Baen Books opens up the Baen Free Library.
- 1999: Webscriptions starts selling unencrypted eBooks.
- 2000: Microsoft launches Microsoft Reader with ClearType technology.
- 2000: Stephen King offers his book "Riding the Bullet" in digital file; it can only be read on a computer.

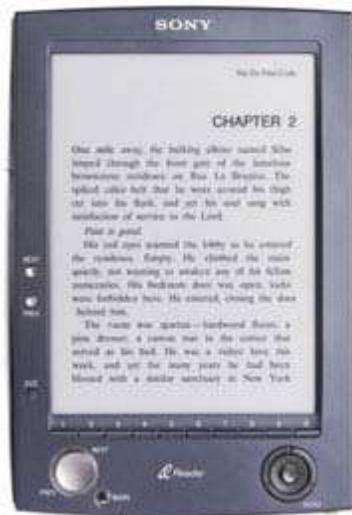
- 2001: Todoebook.com, the first website selling ebooks in Spanish.
- 2002: Random House and HarperCollins start to sell digital versions of their titles in English.



- 2005: Amazon buys Mobipocket.
- 2005: Bookboon.com is launched, allowing people to download free textbooks and travel guide eBooks
- 2006: Sony presents the Sony Reader with e-ink.
- 2006: LibreDigital launched BookBrowse as an online reader for publisher content.
- 2006: BooksOnBoard, the largest independent ebookstore, opens and sells ebooks and audiobooks in six different formats.
- 2007: Zahurk Technologies, Corp, launched the first digital book library on Internet BibliotecaKlemath.com ', loslibrosditaes.com ' and 『digitalbook.us'



- 2007: Amazon launches Kindle in US.
- 2007: Bookeen launched Cybook Gen3 in Europe.
- 2008: Adobe and Sony agreed to share their technologies (Reader and DRM).
- 2008: Sony sells the Sony Reader PRS-505 in UK and France
- 2008: BooksOnBoard is first to sell ebooks for iPhones.
- 2009: Bookeen releases the Cybook Opus in the US and in Europe.
- 2009: Sony releases the Reader Pocket Edition and Reader Touch Edition
- 2009: Amazon releases the Kindle 2.
- 2009: Amazon releases the Kindle DX in the US.
- 2009: Barnes & Noble releases the Nook in the US.
- 2009: Bookboon.com achieves over 10 Million downloads in one year — placing the company as the world's largest publisher of free eBooks
- 2010: Amazon releases the Kindle DX International Edition worldwide.
- 2010: Bookeen reveals the Cybook Orizon at CES.
- 2010: TurboSquid Magazine announces first magazine publication using Apple's iTunes LP format.



- 2010: Apple releases the iPad with an e-book app called iBooks. Since its release in April 2010, as of October, Apple has sold 7 million iPads.
- 2010: Kobo Inc. releases its Kobo eReader to be sold at Indigo/Chapters in Canada and Borders in the United States.
- 2010: Amazon.com reported that its e-book sales outnumbered sales of hardcover books for the first time ever during the second quarter of 2010.
- 2010: Amazon releases the third generation kindle, available in 3G+Wi-Fi and Wi-Fi versions
- 2010: Kobo Inc. releases an updated Kobo eReader which now includes Wi-Fi
- 2010: Barnes & Noble releases the new NOOKcolor

Formats

There are a variety of e-book formats used to create and publish e-books. A writer or publisher has many options when it comes to choosing a format for production. Every format has its proponents and champions, and debates over which format is best can become intense.

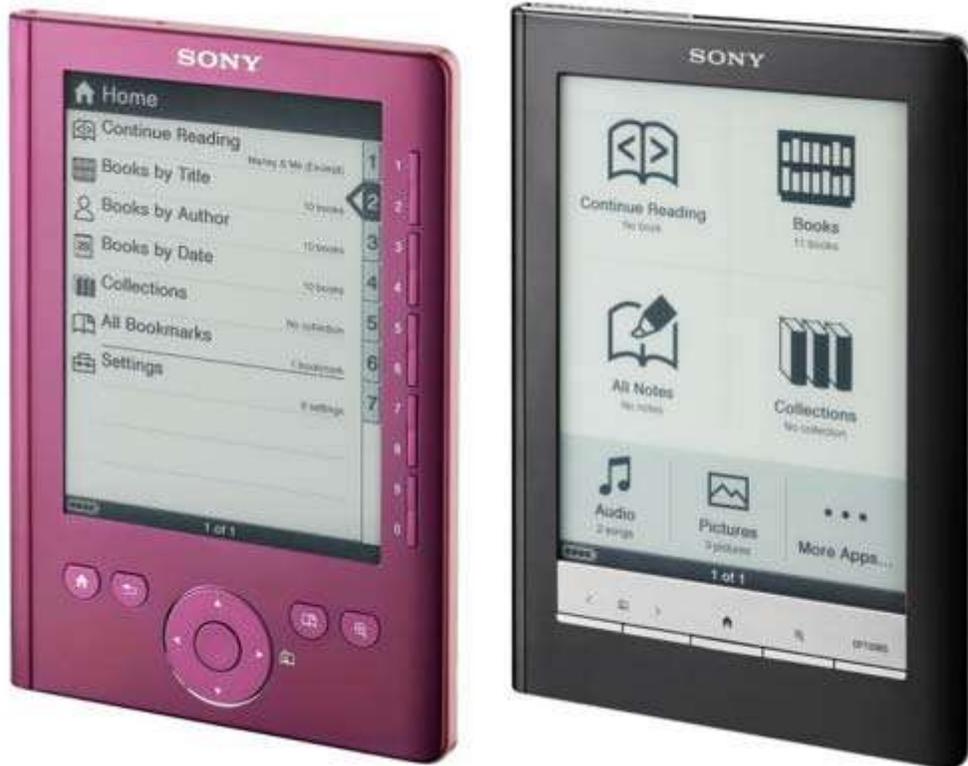
Comparison of e-books with printed books

Advantages

- Availability- There are over 2 million free books available for download as of August 2009. Mobile availability of e-books may be provided for users with a mobile data connection, so that these e-books need not be stored on the device. An e-book can be offered indefinitely, without ever going "out of print".
- Portability and Storage- In the space that a comparably sized print book takes up, an e-reader can potentially contain thousands of e-books, limited only by its

memory capacity. If space is at a premium, such as in a backpack or at home, it can be an advantage that an e-book collection takes up little room or weight.

- Language Accessibility- E-book websites can include the ability to translate books into many different languages, making the works available to speakers of languages not covered by printed translations.



- E-Reader Attributes- Depending on the device, an e-book may be readable in low light or even total darkness. Many newer readers have the ability to display motion, enlarge or change fonts, use Text-to-speech software to read the text aloud, search for key terms, find definitions, or allow highlighting bookmarking and annotation. Devices that utilize E Ink can imitate the look and ease of readability of a printed work while consuming very little power, allowing continuous reading for weeks at time.
 - Costs- While an e-book reader costs much more than one book, the electronic texts are generally cheaper. Moreover, a great share of books are available free of charge. For example, all fiction from before the year 1900 is in the public domain. Also, libraries lend more current e-book titles for limited times, free samples are available of many publications, and there are other lending models being piloted as well. E-books can be printed for less than the price of traditional new books using new on-demand book printers.
 - Security- Depending on possible digital rights management, e-books can be backed up to recover them in the case of loss or damage and it may be possible to recover a new copy without cost from the distributor.

- Distribution- Compared to printed publishing, it is cheaper and easier for authors to self-publish e-books. Also, the dispersal of a free e-book copy can stimulate the sales of the printed version.



- Availability- An e-book can be purchased/borrowed, downloaded, and used immediately, whereas when one buys or borrows a book, one has to go to a bookshop or library during limited hours, or wait for a delivery.
- Environmental Concerns- The production of e-books does not consume paper, ink, etc. Printed books use 3 times more raw materials and 78 times more water to produce albeit they do not require a machine for use

Drawbacks

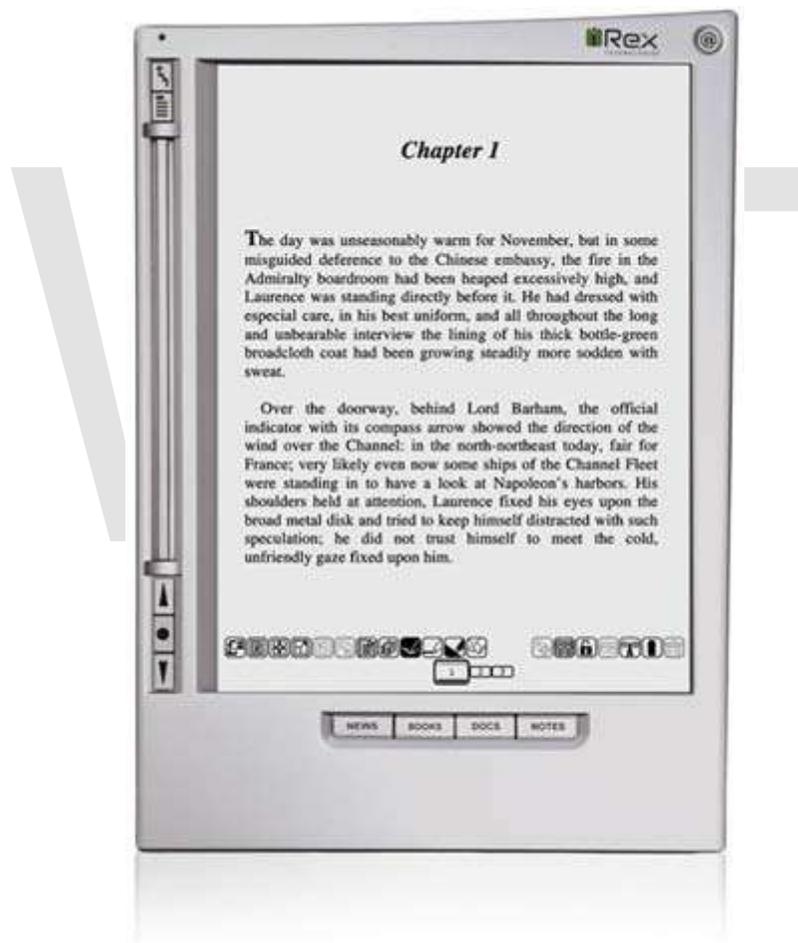
- Changing Technologies- Ebook formats and file types continue to develop and change through time through advances and developments in technology or the introduction of new proprietary formats. While printed books remain readable for many years, e-books may need to be copied or converted to a new carrier or file type over time. PDF and epub are growing standards, but are not universal. The lack of a single universal standard could significantly affect the longevity of some works and their availability or readability in the future as a result of the format(s) used at the time of production.
- Availability of Works- Not all books are available as e-books.
- Aesthetic Appeal- Paper books can be bought and wrapped for a present and a library of books can provide visual appeal, while the digital nature of e-books makes them non-visible or tangible. E-books cannot provide the physical feel of the cover, paper, and binding of the original printed work.
- Power and Shelf Life- A book will never turn off and would be unusable only if damaged or after many decades. The shelf life of a printed book exceeds that of

an e-book reader, as over time the reader's battery will drain and require recharging. Additionally, "As in the case of microfilm, there is no guarantee that [electronic] copies will last. Bits become degraded over time. Documents may get lost in cyberspace...Hardware and software become extinct at a distressing rate."



- Durability- E-book readers are more susceptible to damage from being dropped or hit than a print book. Due to faults in hardware or software, e-book readers may malfunction and data loss can occur. As with any piece of technology, the reader must be protected from the elements (such as extreme cold, heat, water, etc.), while print books are not susceptible to damage from electromagnetic pulses, surges, impacts, or extreme temperatures.
- Artistry and Author's Vision- An author who publishes a book often puts more into the work than simply the words on the pages. E-books may cause people "to do the grazing and quick reading that screens enable, rather than be by themselves with the author's ideas." . They may use the e-books simply for reference purposes rather than reading for pleasure and leisure.
- Costs- The cost of an e-book reader far exceeds that of a single book, and e-books often cost the same as their print versions. Due to the high cost of the initial investment in some form of e-reader, e-books are cost prohibitive to much of the world's population. Furthermore, there is no used e-book market, so consumers will neither be able to recoup some of their costs by selling an unwanted title they have finished, nor will they be able to buy used copies at significant discounts, as they can now easily do with printed books.

- Security- Because of the high-tech appeal of the e-reader, they are a greater target for theft than an individual print book. Along with the theft of the physical device, any e-books it contains also become stolen. E-books purchased from vendors like Amazon or Barnes & Noble.com are stored "in the cloud" on servers and "digital lockers" and have the benefit of being easily retrieved if an e-reading device is lost. Not all e-booksellers are cloud based; if an e-book is stolen, accidentally lost, or deleted, in the absence of a backup it may have to be repurchased.
- Limitations of Readers- The screen resolutions of reading devices are currently lower than actual printed materials. Because of proprietary formats or lack of file support, formatted e-books may be unusable on certain readers. Additionally, the reader's interaction with the reader may cause discomfort, for example glare on the screen or difficulty holding the device.



- Digital Rights Management and Piracy- Due to digital rights management, customers typically cannot resell or loan their e-books to other readers. However, some Barnes & Noble e-books are lendable for two weeks via their 'LendMe' technology. Additionally, the potential for piracy of e-books may make publishers and authors reluctant to distribute digitally.

- Environmental Concerns- E-book readers require various toxic substances to produce, are non-biodegradable, and the disposal of their batteries in particular raises environmental concerns. As technologies rapidly change and old devices become obsolete, there will be larger amounts of toxic wastes that are not easily biodegradable like paper.

From Computer Desktop Encyclopedia
 Reproduced with permission.
 © 1998 NuvolMedia, Inc.



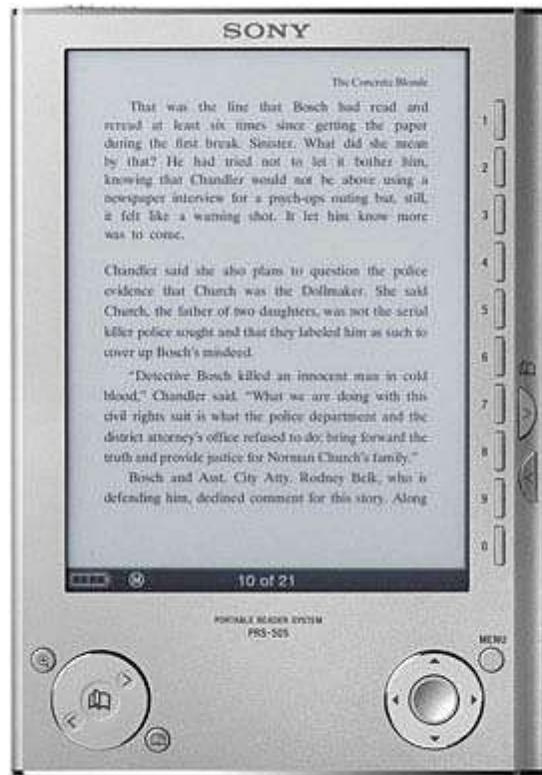
- Privacy Concerns - E books and software can easily track data, times, usage, pages, and details about what one is reading and how often. Similar to this is the growing amount of data available through Google search engines, Facebook, and through data mining. For the first time in history it is now far easier to track and record what specific people might be reading. The notions of privacy, private writing, solitude, and personal reading are changing.
- Picture Books - Books with large pictures (such as children's books) or diagrams are more inconvenient for viewing and reading.



Digital rights management

Anti-circumvention techniques may be used to restrict what the user may do with an e-book. For instance, it may not be possible to transfer ownership of an e-book to another person, though such a transaction is common with physical books. Some devices can phone home to track readers and reading habits, restrict printing, or arbitrarily modify reading material. This includes restricting the copying and distribution of works in the public domain through the use of "click-wrap" licensing, effectively limiting the rights of the public to distribute, sell or use texts in the public domain freely.

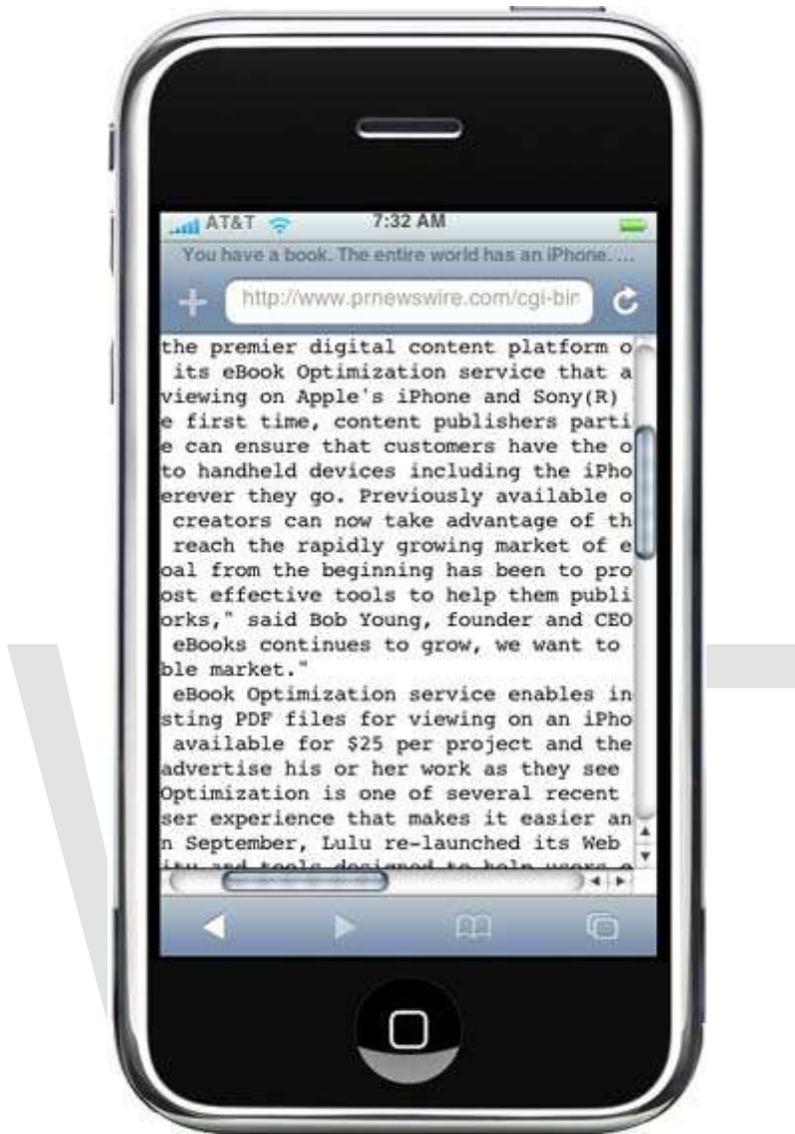
Most e-book publishers do not warn their customers about the possible implications of the digital rights management tied to their products. Generally they claim that digital rights management is meant to prevent copying of the e-book. However in many cases it is also possible that digital rights management will result in the complete denial of access by the purchaser to the e-book. With some formats of DRM, the e-book is tied to a specific computer or device. In these cases the DRM will usually let the purchaser move the book a limited number of times after which he cannot use it on any additional devices. If the purchaser upgrades or replaces their devices eventually they may lose access to their purchase. Some forms of digital rights management depend on the existence of online services to authenticate the purchasers. When the company that provides the service goes out of business or decides to stop providing the service, the purchaser will no longer be able to access the e-book.



As with digital rights management in other media, e-books are more like rental or leasing than purchase. The restricted book comes with a number of restrictions, and eventually access to the purchase can be removed by a number of different parties involved. These include the publisher of the book, the provider of the DRM scheme, and the publisher of the reader software. These are all things that are significantly different from the realm of experiences anyone has had with a physical copy of the book.

Production

Some e-books are produced simultaneously with the production of a printed format, as described in electronic publishing, though in many instances they may not be put on sale until later. Often, e-books are produced from pre-existing hard-copy books, generally by document scanning, sometimes with the use of robotic book scanners, having the technology to quickly scan books without damaging the original print edition. Scanning a book produces a set of image files, which may additionally be converted into text format by an OCR program. Occasionally, as in some e-text projects, a book may be produced by re-entering the text from a keyboard.



As a newer development, sometimes only the electronic version of a book is produced by the publisher. It is also possible to convert electronic book to a printed book by print on demand. However this is an exception as tradition dictates that a book be launched in the print format and later if the author wishes, an electronic version is also produced.

There are some parts of the industry where there are particularly notable leading firms. In the general field of science-fiction and fantasy, Baen Books, an American publishing company established in 1983 by science fiction publishing industry long-timer Jim Baen (1943–2006) has a well-established position. It is a science fiction and fantasy publishing house that specializes in space opera/military science fiction and fantasy (though it does not restrict itself to these subgenres). It is notable for releasing books without DRM in a variety of formats, before hard-copy publication, and pre-releasing ebooks in parts before the hard-copy release. Many older titles are available for free, especially the first book in a series.



As of 2010, there is no industry-wide e-book bestseller list, but various e-book vendors compile bestseller lists, such as those by BooksOnBoard, Amazon Kindle Bestsellers and Fictionwise. There are two yearly awards for excellence in e-books—the EPIC eBook Award (formerly EPPIE) given by EPIC, and the Dream Realm Award for science fiction, fantasy and horror e-books. Both awards have been given since 2000.

e-Readers

e-Readers may be specifically designed for that purpose, or intended for other purposes as well. The term is restricted to hardware devices and used to describe a category type.

Specialized devices have the advantage of doing one thing well. Specifically, they tend to have the right screen size, battery lifespan, lighting and weight. A disadvantage of such devices is that they are often expensive when compared to multi-purpose devices such as laptops and PDAs.

In 2010, competition sent the price for the most popular electronic reading devices below USD 200.

Chapter 2

Comparison of e-Book Formats

The following is a **comparison of e-book formats** used to create and publish e-books.

A writer or publisher has many options when it comes to choosing a format for publication. While the average end-user might arguably simply want to read books, every format has its proponents. The myriad of e-book formats is sometimes referred to as the "Tower of eBabel".

Formats

Formats available include, but are by no means limited to:

Plain text files

Format: text
Published as: .txt

E-books in plain text exist and are very small in size. For example, the Bible, an 800,000-word book, is only about 4 MB. The ASCII standard allows ASCII-only text files (unlike most other file types) to be interchanged and readable on Unix, Macintosh, Microsoft Windows, DOS, and other systems. These differ in their preferred line ending convention and their interpretation of values outside the ASCII range (their character encoding).

Hypertext Markup Language

Format: Hypertext
Published as: .htm; .html

HTML is the markup language used for most web pages. E-books using HTML can be read using a Web browser. The specifications for the format are available without charge from the W3C.

HTML adds specially marked meta-elements to otherwise plain text encoded using character sets like ASCII or UTF-8. As such, suitably formatted files can be, and

sometimes are, generated *by hand* using a *plain text editor* or *programmer's editor*. Many *HTML generator* applications exist to ease this process and often require less intricate knowledge of the format details involved.

HTML on its own is not a particularly efficient format to store information, requiring more storage space for a given work than many other formats. However, several e-Book formats including the Amazon Kindle, Open eBook, Compressed HM, Mobipocket and EPUB use one HTML file for each book chapter and then Zip compress the files, along with images, metadata and style sheets into one file.

HTML files encompass a wide range of standards and displaying HTML files correctly is extremely complicated. Additionally many of the features supported, such as forms, are not relevant to e-books.

Amazon Kindle

Format: Kindle

Published as: .azw

With the launch of the Kindle eBook reader, Amazon.com created the proprietary format, AZW. It is based on the Mobipocket standard, with a slightly different serial number scheme (it uses an asterisk instead of a dollar sign) and its own DRM formatting. Because the eBooks bought on the Kindle are delivered over its wireless system called Whispernet, the user does not see the AZW files during the download process. The Kindle format is now available on a variety of platforms.

Open Electronic Package

Format: Open eBook

Published as: .opf

OPF is an XML-based e-book format created by E-Book Systems.

TomeRaider

Format: TomeRaider

Published as: .tr2; .tr3

The TomeRaider e-book format is a proprietary format. There are versions of TomeRaider for Windows, Windows Mobile (aka Pocket PC), Palm, Symbian, iPhone and more. Several articles are available as TomeRaider files with all articles unabridged, some even with nearly all images. Capabilities of the TomeRaider3 e-book reader vary considerably per platform: the Windows and Windows Mobile editions support full HTML and CSS. The Palm edition supports limited HTML (e.g., no tables, no fonts), and

CSS support is missing. For Symbian there is only the older TomeRaider2 format, which does not render images or offer category search facilities. Despite these differences any TomeRaider e-book can be browsed on all supported platforms. The Tomeraider website claims to have over 4000 e-books available, including free versions of the Internet Movie Database.

Arghos Diffusion

Format: Arghos Reader

Published as: .aeh

The AEH format is an XML-based proprietary format developed by the French firm Arghos Diffusion. AEH files use a proprietary DRM and encryption method and are readable only in the *Arghos Player*. It supports various input formats for text, audio or video, such as PDF, WMA, MP3, WMV, and allows multiple interactive functions such as bookmarking, advanced plain-text searching, dynamic text highlighting, etc.

Flip Books

Format: Interaxive media

Published as:

A "Flip Book" is a type of E-Book distinguished by virtual pages that actually "flip", much like turning pages of paper in a real book or magazine. The first dynamic Flip Book Reader was developed in 2003/2004 by Interaxive Media for Nishe Media (Canada) and was therefore called "Nishe Pages". The first version was produced in part by Cybaris (Canada) and was first publicly showcased in August 2004. Soon thereafter, many copycat "flip books" started appearing thanks to technological advances in Macromedia Flash, mostly hard coded using Flash components.

The original software remains unique in that it is powered by a complete server-based CMS system that allows the books to be created, published, and viewed remotely from a web server without requiring any custom software to be installed. Nishe Media went defunct in 2004, leaving the unfinished software to Interaxive Media who continued its development in Hong Kong. Though not widely used outside of Asia, it is now at version 3.0 and can be a server-based E-Book platform. It remains privately held by the original developer, Ryan Sutherland, owner and founder of Interaxive Media.

ANSI/NISO Z39.86 (DAISY)

Format: DAISY

Published as:

The Digital Accessible Information SYstem (DAISY) is an XML-based open standard maintained by the DAISY Consortium for people with print disabilities. DAISY has wide international support with features for multimedia, navigation and synchronization. A subset of the DAISY format has been adopted by law in the United States as the National Instructional Material Accessibility Standard (NIMAS), and K-12 textbooks and instructional materials are now required to be provided to students with disabilities.

DAISY is already aligned with the EPUB open standard, and is expected to fully converge with its forthcoming EPUB3 revision.

FictionBook (Fb2)

Format: FictionBook

Published as: .fb2

FictionBook is a popular XML-based e-book format, supported by free readers such as FBReader, Haali Reader and STDU Viewer.

Text Encoding Initiative

Format: TEI Lite

Published as: .xml

TEI Lite is the most popular of the TEI-based (and thus XML-based or SGML-based) electronic text formats.

Plucker

Format: Plucker

Published as:

Plucker is a free e-book reader application with its own associated file format and software to automatically generate plucker files from HTML files, web sites or RSS feeds. The format is a compressed HTML archive, somewhat like Microsoft's CHM.

Compressed HM

Format: Microsoft Compressed HTML Help

Published as: .chm

CHM format is a proprietary format based on HTML. Multiple pages and embedded graphics are distributed along with proprietary metadata as a single compressed file. In contrast, in HTML, a site consists of multiple HTML files and associated image files in standardized formats.

Portable Document Format

Format: Adobe Portable Document Format

Published as: .pdf

A file format created by Adobe Systems, initially to provide a standard form for storing and editing printed publishable documents. The format derives from PostScript, but without language features like loops, and with added support for features like compression and passwords. Because PDF documents can easily be viewed and printed by users on a variety of computer platforms, they are very common on the World Wide Web. The specification of the format is available without charge from Adobe.

PDF files typically contain brochures, product manuals, magazine articles — up to entire books, as they can embed fonts, images, and other documents. A PDF file contains one or more zoomable page images.

Since the format is designed to reproduce page images, the text traditionally could not be re-flowed to fit the screen width or size. As a result PDF files designed for printing on standard paper sizes are less easily viewed on screens with limited size or resolution, such as those found on mobile phones and PDAs. Adobe has addressed this by adding a re-flow facility to its Acrobat Reader software, but for this to work the document must be marked for re-flowing at creation, which means that existing PDF documents will not benefit unless they are tagged and resaved. The Windows Mobile (aka Pocket PC) version of Adobe Acrobat will automatically attempt to tag a PDF for reflow during the synchronization process using an installed plugin to Active Sync. However, this tagging process will not work on most locked or password protected PDF documents. It also doesn't work at present (2009–10) on the Windows Mobile Device Center (the successor to Active Sync) as found in Windows Vista and Windows 7. This limits automatic tagging support during synchronization to Windows XP/2000.

Multiple products support creating and tagging PDF files, such as Adobe Acrobat, PDFCreator, OpenOffice.org, iText, and FOP, and several programming libraries. Adobe Reader (formerly called *Acrobat Reader*) is Adobe's product used to view PDF files; third party viewers such as xpdf are also available. Mac OS X has built-in PDF support, both for creation as part of the printing system and for display using the built-in Preview application.

Later versions of the specification add support for forms, comments, hypertext links, and even interactive elements such as buttons for forms entry and for triggering sound and video. Such features may not be supported by older or third-party viewers and some are not transferable to print.

PDF files are supported on the following e-book readers: Mobipocket, iRex iLiad, iRex DR1000, Sony Reader, Bookeen Cybook, Foxit eSlick, Amazon Kindle (1, 2, International & DX), Barnes & Noble Nook, the iPad, PocketBook Reader, Bebook Neo and the Kobo eReader.

PostScript

Format: PostScript

Published as: ps

PostScript is a page description language used in the electronic and desktop publishing areas **for defining** the contents and layout of a printed page, which can be used by a rendering program to assemble and create the actual output bitmap. Many office printers directly support interpreting PostScript and printing the result. As a result, the format also sees wide use in the Unix world.

DjVu

Format: DjVu

Published as: .djvu

DjVu is a format that specializes in and particularly excels at storing scanned images. It includes advanced compressors optimized for low-color images, such as text documents. Individual files may contain one or more pages. DjVu files cannot be re-flowed.

The contained page images are divided in separate layers (such as multi-color, low-resolution, background layer using lossy compression, and few-colors, high-resolution, tightly-compressed foreground layer), each compressed in the best available method. The format is designed to decompress very quickly, even faster than vector-based formats.

The advantage of DjVu is that it is possible to take a high-resolution scan (300-400 DPI), good enough for both on-screen reading and printing, and store it very efficiently. Several dozens of 300 DPI black-and-white scans can be stored in less than a megabyte.

Microsoft LIT

Format: Microsoft Reader

Published as: .lit

DRM-protected LIT files are only readable in the proprietary Microsoft Reader program, as the .LIT format, otherwise similar to Microsoft's CHM format, includes Digital Rights Management features. Other third party readers, such as Lexcycle Stanza, can read unprotected LIT files. There are also tools such as Convert Lit, which can convert .lit files to HTML files or OEBPS files.

The Microsoft Reader uses patented ClearType display technology. In Reader navigation works with a keyboard, mouse, stylus, or through electronic bookmarks. The Catalog Library records reader books in a personalized "home page", and books are displayed with ClearType to improve readability. A user can add annotations and notes to any page,

create large-print e-books with a single command, or create free-form drawings on the reader pages. A built-in dictionary allows the user to look up words.

eReader

Formerly Palm Digital Media/Peanut Press

Format: Palm Media

Published as: .pdb

eReader is a freeware program for viewing Palm Digital Media electronic books. Versions are available for iPhone, PalmOS, Android, Symbian, BlackBerry, Windows Mobile Pocket PC/Smartphone, desktop Windows, and Macintosh. The reader shows text one page at a time, as paper books do. eReader supports embedded hyperlinks and images. Additionally, the Stanza application for the iPhone and iPod Touch can read both encrypted and unencrypted eReader files.

The company's web site - ereader.com maintains a wide selection of eReader-formatted e-books, available for purchase and download, with a handful of public domain titles available for free. Those books that aren't free are encrypted, with the key being the purchaser's full name and credit card number. This information is not preserved in the e-book. A one-way hash is used, so there no risk of the user's information being extracted.

The program supports features like bookmarks and footnotes, enabling the user to mark any page with a bookmark, and any part of the text with a footnote-like commentary. Footnotes can later be exported as a Memo document.

The company also offers two Windows/MacOS programs for producing e-books: the Dropbook, which is free, and the eBook Studio, which is not. Dropbook is a file-oriented PML-to-PDB converter; eBook Studio incorporates a WYSIWYG editor. Both programs are compatible with simple text files.

There is also support for an integrated reference dictionary (with many options up to and including a 476,000-word Merriam-Webster Dictionary, including pronunciation keys) so that any word in the text can be highlighted and looked up on the dictionary instantly. Commercial fonts can also be individually purchased and downloaded at the company's web site, ereader.com.

On July 20, 2009, Barnes & Noble announced that the eReader format will be the method they will use to deliver e-books. Updated versions of the Palm Digital programs for Apple iPhone/Touch, Blackberry, Mac OS X, and Windows platforms were made available on the Barnes & Noble eBooks website.

On October 20, 2009, Barnes & Noble announced that their Nook Reader will support the eReader format. eReader format is also supported by the discontinued eSlick, an e-reading device from Foxit Software.

Desktop Author

Format: DNL Reader

Published as: .dnl; .exe

Desktop Author is an electronic publishing suite that allows creation of digital web books with virtual turning pages. Digital web books of any publication type can be written in this format, including brochures, e-books, digital photo albums, e-cards, digital diaries, online resumes, quizzes, exams, tests, forms and surveys. DesktopAuthor packages the e-book into a ".dnl" or ".exe" book. Each can be a single, plain stand-alone executable file which does not require any other programs to view it. DNL files can be viewed inside a web browser or stand-alone via the *DNL Reader*.

DNL format is an e-Book format, one which replicates the real life alternative, namely page turning Books. The DNL e-Book is developed by DNAML Pty Limited an Australian company established in 1999. A DNL e-Book can be produced using DeskTop Author or DeskTop Communicator.

Newton eBook

Format: Newton eBook

Published as: .pkg

Commonly known as an Apple Newton book; a single Newton package file can contain multiple books (for example, the three books of a trilogy might be packaged together). All systems running the Newton operating system (the most common include the Newton MessagePads, eMates, Siemens Secretary Stations, Motorola Marcos, Digital Ocean Seahorses and Tarpons) have built-in support for viewing Newton books. The Newton package format was released to the public by Newton, Inc. prior to that company's absorption into Apple Computer. The format is thus arguably open and various people have written readers for it (writing a Newton book converter has even been assigned as a university-level class project).

Newton books have no support for DRM or encryption. They do support internal links, potentially multiple tables of contents and indexes, embedded gray scale images, and even some scripting capability (for example, it's possible to make a book in which the reader can influence the outcome). Newton books utilize Unicode and are thus available in numerous languages. An individual Newton book may actually contain multiple views representing the same content in different ways (such as for different screen resolutions).

Founder Electronics

Format: Apabi Reader

Published as: .xeb; .ceb

APABI is a format devised by Founder Electronics. It is a popular format for Chinese e-books. It can be read using the Apabi Reader software, and produced using Apabi Publisher. Both .xeb and .ceb files are encoded binary files. The Iliad e-book device includes an Apabi 'viewer'.

Libris

Format: Mobile Information Device Profile

Published as: .lbr; .bin

Libris is a Java based eBook reader for mobile devices such as cell phones. Libris will run on most Java enabled devices that support MIDP. The reader formats books to fit the device screen, and shows one page at a time using high quality anti-aliased fonts. Books may employ encryption or be unrestricted. Libris content may be produced using the MakeLibris tool. The Libris reader also supports the PalmDoc format.

Mobipocket

Format: Mobipocket

Published as: .prc; .mobi

The Mobipocket e-book format based on the Open eBook standard using XHTML and can include JavaScript and frames. It also supports native SQL queries to be used with embedded databases. There is a corresponding e-book reader.

The Mobipocket Reader has a home page library. Readers can add blank pages in any part of a book and add free-hand drawings. Annotations — highlights, bookmarks, corrections, notes, and drawings — can be applied, organized, and recalled from a single location. Images are converted to GIF format and have a maximum size of 64K, sufficient for mobile phones with small screens, but rather restrictive for newer gadgets. Mobipocket Reader has electronic bookmarks, and a built-in dictionary.

The reader has a full screen mode for reading and support for many PDAs, Communicators, and Smartphones. Mobipocket products support most Windows, Symbian, BlackBerry and Palm operating systems. Using WINE, the reader works under Linux or Mac OS X. Third-party applications like Okular and FBReader can also be used under Linux or Mac OS X, but they work only with unencrypted files.

The Amazon Kindle's AZW format is basically just the Mobipocket format with a slightly different serial number scheme (it uses an asterisk instead of a Dollar sign), and .prc publications can be read directly on the Kindle.

Mobipocket has developed an .epub to .mobi converter called KindleGen (supports IDPF 1.0 and IDPF 2.0 epub format, according to the company).

Notably, Eastern European letters with diacritical marks are not supported.

EPUB

Format: IDPF/EPUB

Published as: .epub



ePUB

The EPUB logo.

The .epub or OEBPS format is an open standard for e-books created by the International Digital Publishing Forum (IDPF). It combines three IDPF open standards:

- Open Publication Structure (OPS) 2.0, which describes the content markup (either XHTML or Daisy DTBook)
- Open Packaging Format (OPF) 2.0, which describes the structure of an .epub in XML
- OEBPS Container Format (OCF) 1.0, which bundles files together (as a renamed ZIP file)

Currently, the format can be read by the Kobo eReader, Apple iBooks, Barnes and Noble Nook, Sony Reader, BeBook, Bookeen Cybook Gen3 (with firmware v. 2 and up), COOL-ER, Adobe Digital Editions, Lexcycle Stanza, BookGlutton, AZARDI, Aldiko and WordPlayer on Android and the Mozilla Firefox add-on EPUBReader. Several other reader software programs are currently implementing support for the format, such as dotReader, FBReader, Mobipocket, uBook and Okular. Another software .epub reader, Lucidor, is in beta.

Adobe Digital Edition uses .epub format for its e-books, with DRM protection provided through their proprietary ADEPT mechanism. The recently developed INEPT framework and scripts have been reverse-engineered to circumvent this DRM system.

DSLlibris, a Sourceforge.net project, is able to decode e-books in .epub and .xht format for reading on Nintendo DS systems.

Broadband eBooks (BBeB)

Format: Sony media

Published as: .lrf; .lrx

The digital book format used by Sony Corporation. It is a proprietary format, but some reader software for general-purpose computers, particularly under Linux (for example, calibre's internal viewer), has the capability to read it. The LRX file extension represents a DRM encrypted eBook.

SSReader

Format: SSReader

Published as: .pdg

The digital book format used by a popular digital library company 超星数字图书馆 in China. It is a proprietary raster image compression and binding format, with reading time OCR plug-in modules. The company scanned a huge number of Chinese books in the China National Library and this becomes the major stock of their service. The detailed format is not published. There are also some other commercial e-book formats used in Chinese digital libraries.

TealDoc

Format: TealDoc

Published as: .pdb

TealPoint Software's proprietary reader for Palm OS. In addition to its own format, it opens plain text and PalmDoc files. Newer versions of the software include an editor for Palm OS. Embedded images must be converted to TealPoint's proprietary TealPaint format. The format uses HTML like tags for formatting and has been reverse-engineered for 3rd party programs to edit and convert to/from TealDoc format.

IEC 62448

Format: IEC 62448

Published as:

IEC 62448 is an international standard created by International Electrotechnical Commission (IEC), Technical Committee 100, Technical Area 10 (Multimedia e-publishing and e-book).

The current version of IEC 62448 is an umbrella standard that contains as appendices two concrete formats, XMDF of Sharp and BBeB of Sony. However, BBeB has been discontinued by Sony and the version of XMDF that is in the specification is out of date. The IEC TA10 group is discussing next steps, and has invited the IDPF organization which has standardized EPUB to be a liaison. It is possible that the current version of EPUB and/or the forthcoming EPUB3 revision may be added to IEC 62448. Meanwhile a number of Japanese companies have proposed that IEC standardize a proposed new Japanese-centric file format that is expected to unify DotBook of Voyager Japan and XMDF of Sharp. This new format has not been publicly disclosed as of November, 2010 but it is supposed to cover basic representations for the Japanese language. Technically speaking, this revision is supposed to provide a Japanese minimum set, a Japanese extension set, and a stylesheet language. These issues were discussed in the TC100 meeting held in October 2010 but no decisions were taken besides offering the liaison status to IDPF.

Comic Book Archive file

<i>Format:</i>	compressed images
<i>Published as:</i>	.cbr → RAR .cbz → ZIP .cb7 → 7z .cbt → TAR .cba → ACE

A Comic Book Archive file or ComicBook Reader File consists of a series of image files, typically PNG (lossless compression) or JPEG (lossy compression) files, stored as a single archive file, for the purpose of sequential viewing of images, especially comic books. The idea was made popular by the CDisplay image viewer; since then, many viewers for different platforms have been created. Comic Book Archive files are not a distinct file format; only the file name extension differs from a standard file of the given archive type. Some applications support additional tag information (like artists or story information) in the form of embedded XML files in the archive, or use of the Zip comment function.

Multimedia eBooks

<i>Format:</i>	Eveda
<i>Published as:</i>	.exe or .html

A multimedia ebook is media and book content that utilizes a combination of different book content formats. The term can be used as a noun (a medium with multiple content formats) or as an adjective describing a medium as having multiple content formats.

The 'multimedia ebook' term is used in contrast to media which only utilize traditional forms of printed or text books. Multimedia ebooks include a combination of text, audio, images, video, and/or interactive content formats. Much like how a traditional book can contain images to help the text tell a story, a multimedia ebook can contain other elements not formerly possible to help tell the story.

With the advent of more widespread tablet-like computers, such as the smartphone, some publishing houses are planning to make multimedia ebooks, such as Penguin.

Features and hardware tables

Features

Format	File name extension	DRM support	Image support	Table support	Sound support	Interactivity support	Word wrap support	Open standard	Embedded annotation support	Book-marking
Plain text	.txt	No	No	No	No	No	Yes	Yes	No	No
HTML	.html	No	Yes	Yes	No	No	Yes	Yes	No	No
PostScript	.ps	No	Yes	?	No	No	No	Yes	?	?
Portable Document Format	.pdf	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
DjVu	.djvu	?	Yes	Yes	No	No	No	Yes	Yes	Yes
EPUB (IDPF)	.epub	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FictionBook	.fb2	Yes	Yes	?	No	No	Yes	Yes	Yes	?
Mobipocket	.prc, .mobi	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Kindle eReader	.azw, .pdb	Yes	Yes	No	No	No	Yes	No	Yes	Yes
TealDoc	.pdb	Yes	Yes	?	No	No	Yes	Yes	?	Yes
Broadband eBook	.lrf, .lrx	Yes	Yes	?	No	No	Yes	No	?	?
WOLF[disambiguation needed]	.wol	Yes	Yes	?	No	No	No	No	?	?
Tome Raider	.tr2, .tr3	Yes	Yes	?	No	No	Yes	No	?	?
ArgghosReader	.aeh	Yes	Yes	?	No	No	Yes	No	?	Yes
Microsoft Reader	.lit	Yes	Yes	?	No	No	Yes	No	?	Yes
Multimedia EBook	.exe	Yes	Yes	?	Yes	Yes	No	Yes	Yes	Yes
Repligo	.rgo	?	Yes	Yes	No	No	Yes	No	No	No

¹ Proprietary format - ² Predecessor of ePUB - ³ Versions support either ePUB or MobiPocket - ⁴ Only ePUB version and with FW 2.0+ - ⁵ Requires additional software

WWT

Chapter 3

ILiad

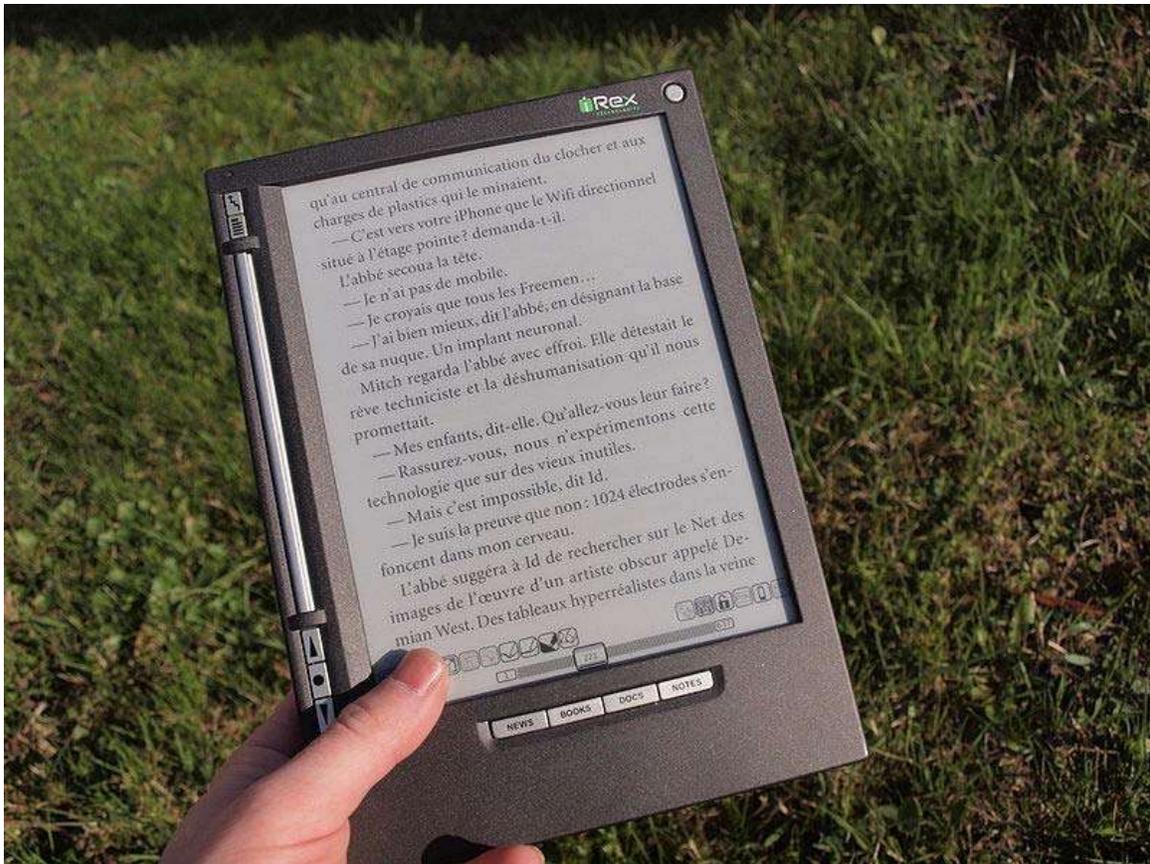
iRex iLiad



Manufacturer	iRex Technologies
Screen	124×152 mm (W×H), 768×1024 pixels, 160 ppi density, 16-level grayscale Electronic paper.
Operating system	Linux (2.4 kernel)
Input	Wacom touchscreen, Graphical user interface, next/prev/quick access buttons.

CPU	400MHz Intel XScale.
Memory	64 MB RAM, 256 MB (128 MB available) internal storage, Expandable via USB, MMC or CF cards.
Networks	MyiRex account WiFi 802.11g,
Connectivity	10/100 Mbit/s Ethernet LAN (via travel hub), USB port, audio jack.
Battery	Lithium Ion battery.
Physical size	155x217x16 mm (WxHxD)
Weight	389 g (13.7 oz)

The **iLiad** is an electronic handheld device, or e-Reader, which can be used for document reading and editing. Like the Barnes and Noble nook, Sony Reader or Amazon Kindle, the iLiad makes use of an electronic paper display.



iLiad in sunlight



iLiad e-book reader equipped with e-paper display



Description

Main specifications:

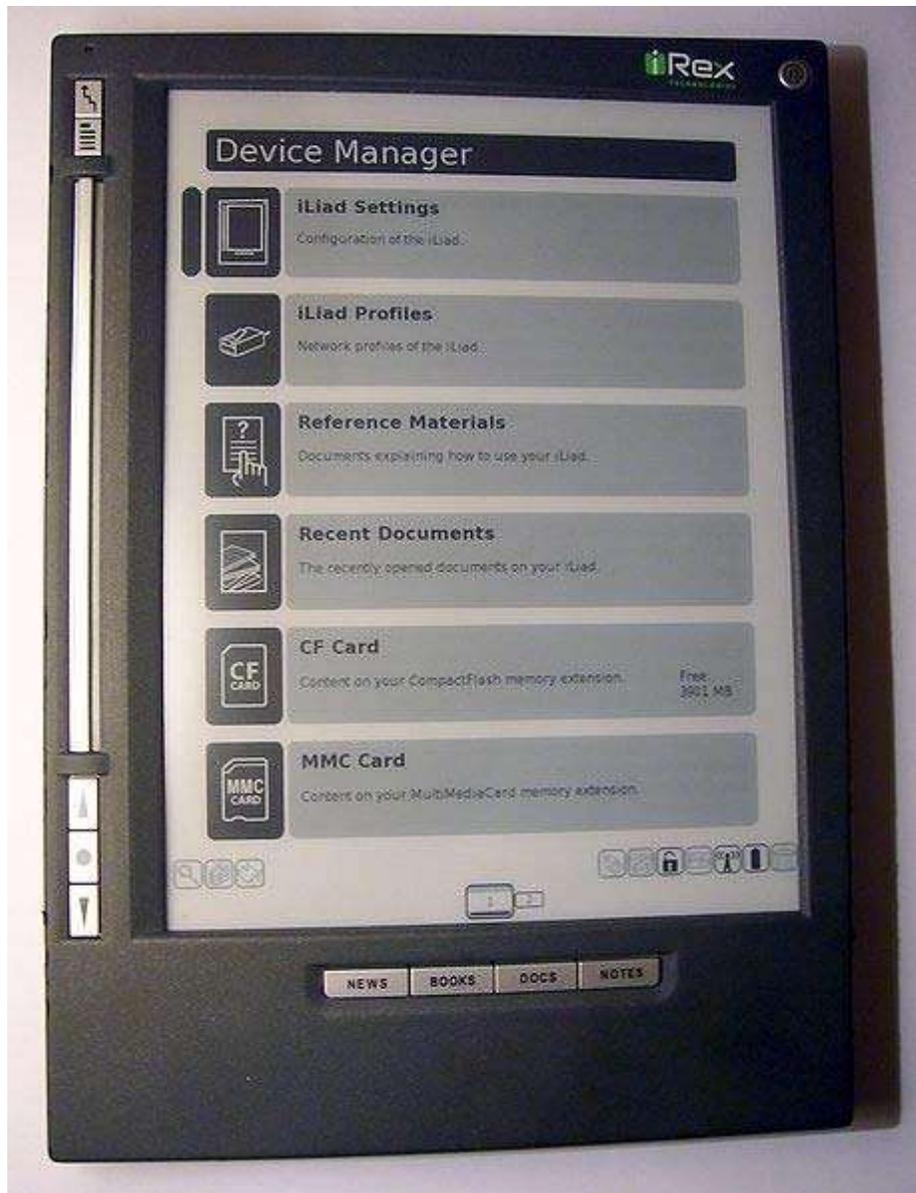
- an 8.1-inch (21 cm) electronic paper display, area for displaying content is 124x165mm
- a resolution of 768x1024 pixels (160 dpi)
- 16 levels of grayscale
- a USB connector for external storage,
- a CompactFlash Type II slot for memory extension or other applications
- a MultiMediaCard slot for MMC memory cards
- a 3.5 mm stereo audio jack for a headset

- WiFi 802.11g wireless LAN
- 10/100 Mbit/s wired LAN
- 390 grams (14 oz) weight
- 400 MHz Intel XScale processor
- 64MB RAM
- 256MB internal flash memory (128 for user, 128 for system)
- Linux-based operating system (2.4 kernel)
- SDK is provided , so functionality is easily extended.

It measures 155 mm × 216 mm × 16 mm (width × height × depth), the size of an A5 document, or roughly a 6"×9" steno notebook. The display used is an active matrix electrophoretic display, which uses E Ink Vizplex Imaging Film manufactured by E Ink Corporation. Underneath the E Ink screen is a digitizing tablet by Wacom which requires a stylus for input. When it was introduced, the Iliad had largest screen size of existing e-paper products, but the newer iRex Digital Reader 1000's 10.2-inch (26 cm) display is now the largest in production.

The iLiad is capable of displaying document files in a number of formats, including PDF, Mobipocket, XHTML and plain text. It can also display JPEG, BMP and PNG images, but not in color. As of May 3, 2007 Mobipocket is supported, making the mobipocket digital rights management (DRM) content available on this platform . iRex's product page for the iLiad states that "Support for additional E-book formats will become available over the coming months."

Through its wireless service, iDS, the iLiad can also directly download content. Les Echos, a French financial newspaper, is distributed this way, as is Dutch newspaper NRC Handelsblad, and additional content is expected. Users can connect to their computer over a wireless network to sync new data onto the iLiad's internal memory or an inserted MMC, SD, or CF card. More wireless and network functions are to be delivered in upcoming firmware updates.



The distributor of the iLiad is iRex Technologies, a Philips spin-off company. It was initially advertised in December 2005, to be launched in April 2006, but was delayed until July, when it started to be sold as a beta product. It was released to the general public near the end of July, and since then has undergone considerable software revisions.

Its list price in Europe is €649, and in US \$699, however it is no longer available in North America due to FCC regulation non-compliance.

Advanced features

One of the advanced features of the iLiad is the ability to add notes to existing documents. With the integrated Wacom tablet and stylus, it is possible to write directly on almost any document and those notes will remain on that document whenever it is viewed on the iLiad. Using the desktop software, those notes can be merged into the original document. This provides malleability, an important feature of physical books that is missing from most ebook products, allowing users to annotate, highlight, and personalize the text.

Third-party development

Because of its open Linux operating system, the iLiad is able to run third party applications created for it. Developers and users wishing to create or run third party applications can request shell access from the manufacturer.

Developers have been able to improve on the device's functionality by porting viewers such as FBReader, and programs such as abiword and stardict. Full screen PDF reading is made available by community-supported iPDF releases. Programs for recreation, including audio playback, sudoku, and calendars, are rapidly growing community content available for use on the iLiad. Independent users have also reported successful porting of mobile web browsers to iLiad's Linux platform, although with limited functionality and many bugs.

iRex as a company has had a lukewarm relationship with its Open Source developers. Most of the complaints toward iRex center around speed of release of SDKs and other information. They have, however, also been quite forthcoming on some projects and some of the changes the community has made have been integrated back into the company's software distribution, perhaps the most notable being that of stylus and input calibration, an external development project led by Jay Kuri and published as part of the main distribution in early 2008.

Version 2

In September 2007, iRex Technologies released an update to the iLiad. While officially called "iLiad 2nd Edition", it is generally noted to be a minor update to the original.

The update includes:

- Redesigned backplane
- Increased battery capacity
- Software version 2.11
- Updated travel charger
- Included case

The 2.11 software, which contains stylus calibration, extended battery life, and other things, is also available to first-generation iLiads.

Book edition

In May 2008, iRex Technologies added a third installment to the iLiad line of products, this time branded under the name iLiad Book Edition. This is the iLiad Version 2 without WiFi and a new silver look. The technical cutbacks reduce the price to \$599 (€499), which is cheaper than the original. It also comes with 50 free classics, including works from well-known writers Jules Verne, Charles Dickens, Lewis Carroll, and Leo Tolstoy.

Bankruptcy

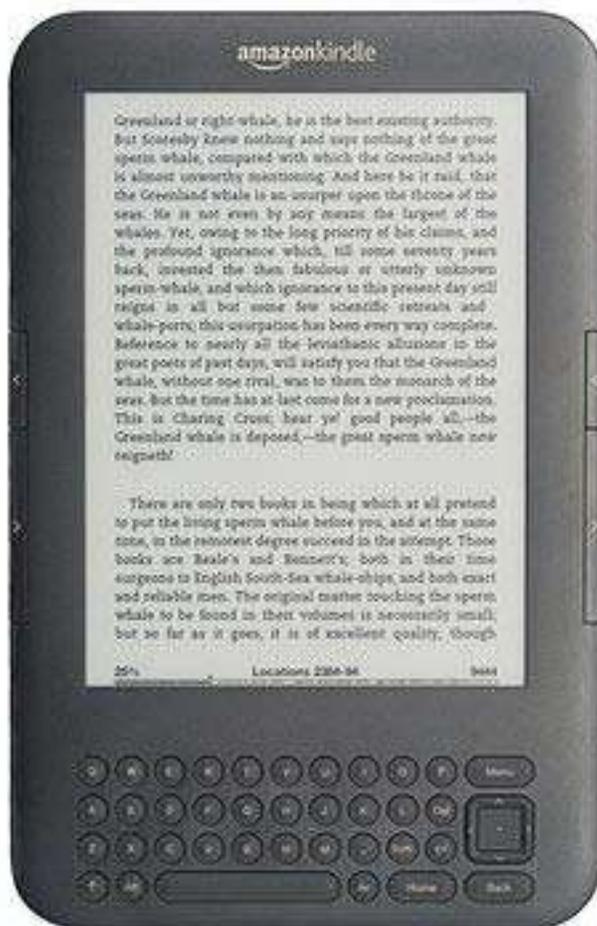
In early June 2010, iRex Technologies of the Netherlands filed for Bankruptcy protection, after the FCC blocked sales of their DR800 reader in the US market.

WWT

Chapter 4

Amazon Kindle

Amazon Kindle



A third generation Kindle in graphite color

Manufacturer Foxconn for Amazon.com

Type E-book reader

Release date 1st generation: November 19, 2007

2nd generation: February 9, 2009
3rd Generation: August 27, 2010
Kindle DX (1st Gen): June 10, 2009
Kindle DX (2nd Gen): Jan 19, 2010
Kindle DX Graphite (3rd Gen): July 7,
2010

**Operating
system**

Linux-2.6.10

Power

3.7 V, 1530 mAh lithium polymer,
BA1001
model

CPU

Freescale 532 MHz, ARM-11

**Storage
capacity**

(total/user available) 256 MB/180 MB
(original) or 2 GB/1.4 GB (Kindle 2) or
4 GB/3 GB (Kindle 3) or 4 GB/3.3 GB
(Kindle DX) internal flash memory

Display

6 in diagonal,
3.6 in (91 mm) × 4.8 in (122 mm),
600 × 800 pixels or 0.48 megapixels,
167 ppi density,
4-level grayscale (original)
or 16-level grayscale (Kindle 2 and 3)
electronic paper

Input

USB 2.0 port (micro-B connector),
SD card (original model only),
3.5 mm stereo headphone jack
built-in speaker,
AC power adapter jack

Connectivity

Amazon Whispernet using EVDO/CDMA
AnyDATA wireless modem (selected
models),
802.11bg WiFi (Kindle 3)

	8.0 × 5.3 × 0.8 in (203 × 135 × 20.3 mm)
	(original)
	8.0 × 5.3 × 0.36 in (203 × 135 × 9.14 mm)
Dimensions	(Kindle 2)
	7.5 × 4.8 × 0.34 in (190 × 123 × 8.51 mm)
	(Kindle 3)
	10.4 x 7.2 x 0.38 in (Kindle DX 2)
	10.2 oz (290 g) (Kindle 1 & 2)
Weight	8.7 oz (247 g) (Kindle 3)
	8.5 oz (241 g) (Kindle 3 Wi-Fi only)
	18.9 oz (540 g) (Kindle DX 2)

Amazon Kindle is a portable e-book reader. More precisely, it is a software, hardware and network platform developed by Amazon.com (subsidiary Lab126) that utilizes wireless connectivity to enable users to shop for, download, browse, and read e-books, newspapers, magazines, blogs, and other digital media in some countries.

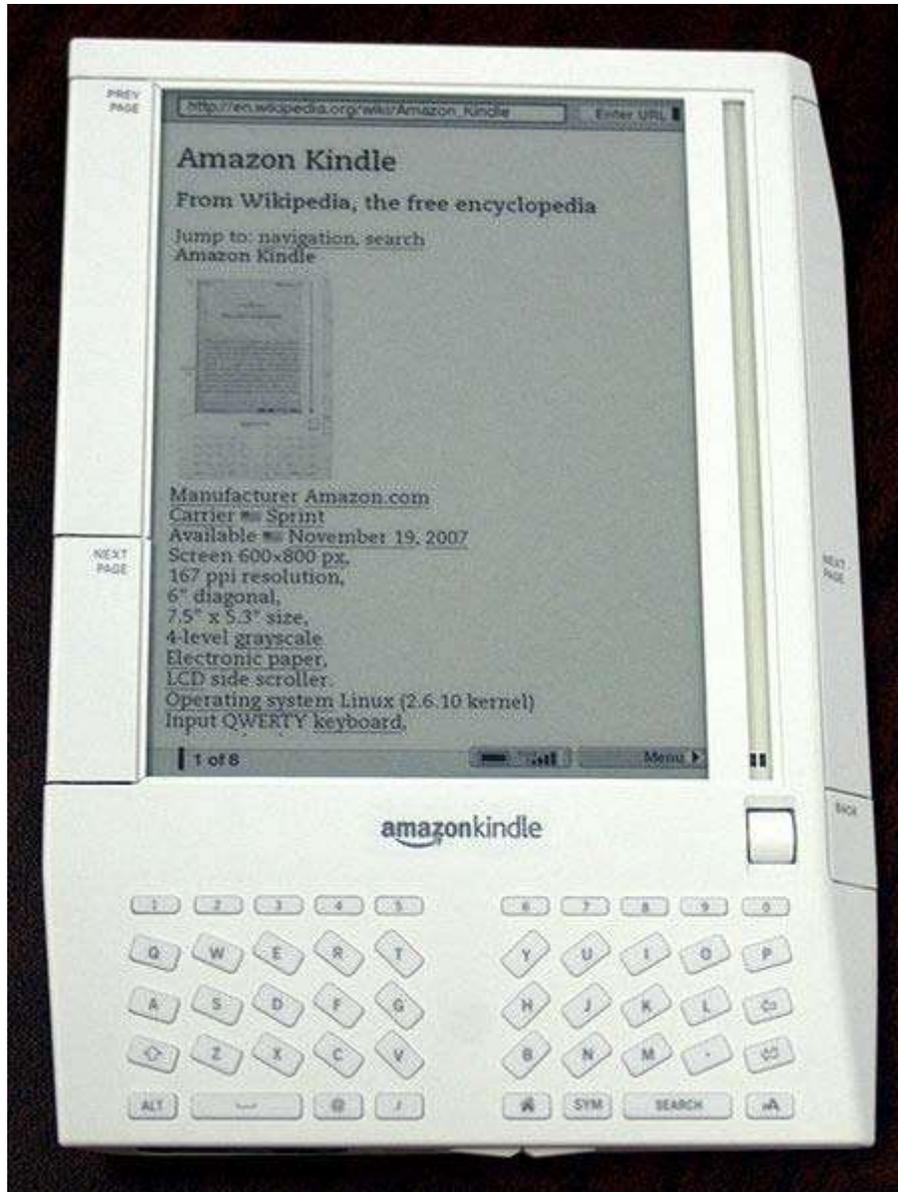
Several hardware devices support this platform, including a main "Kindle" line and a parallel "Kindle DX" line. The most recent Kindle device is the third incarnation of the main line, officially named "Kindle", but usually referred to as "Kindle 3". Kindle 3 was released on August 27, 2010. User reports indicate that the new display, with E Ink Pearl technology on the Kindle 3 is noticeably superior to previous generations.

The Kindle DX line features larger screens than the main Kindle line. It is marketed as more suitable for displaying newspaper and textbook content, however, no Kindle model preserves the print edition page numbers used when citing sources in school papers and scholarly research. Amazon has also introduced Kindle software for use on various devices and platforms, including Microsoft Windows, iOS, BlackBerry, Mac OS X (10.5 onwards), Android and soon Windows Phone 7.

The Kindle hardware devices use an e-ink electronic paper display that features 16 shades of grey. This allows for a long battery life and easy readability. Content for the Kindle can be purchased online and downloaded wirelessly in some countries, using either standard Wi-Fi or Amazon's 3G "Whispernet" network. Whispernet is accessible without any monthly fee or wireless subscription, although fees can be incurred for the delivery of periodicals and other content when roaming internationally beyond the customer's home country. Through a service called "Whispersync", customers can synchronize reading progress, bookmarks and other information across Kindle hardware devices and other mobile devices.

History

Original version



A first generation Kindle

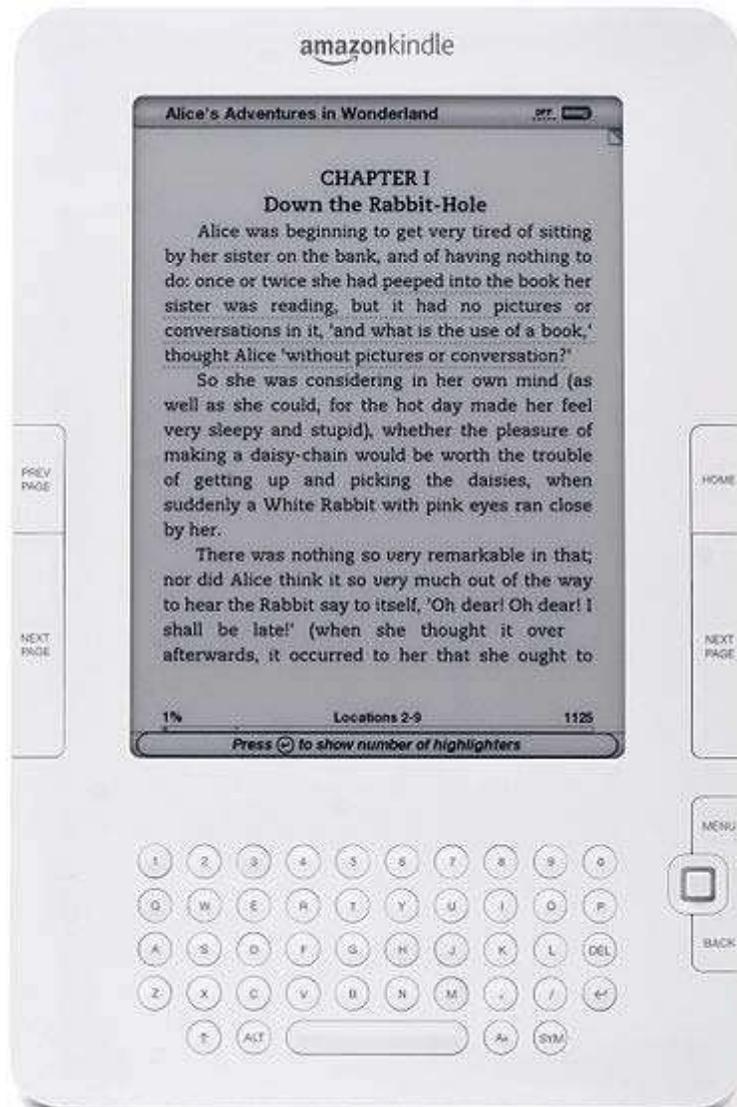
Amazon released the Kindle First Generation on November 19, 2007, for US\$399 and it sold out in five and a half hours. The device remained out of stock for five months until late April 2008.

It is the only Kindle with expandable memory, via an SD card slot.

The device features a 6 inch (diagonal) 4-level grayscale display, with 250 MB of internal memory, which can hold approximately 200 non-illustrated titles.

Amazon did not sell the Kindle First Generation outside the United States. Plans for a launch in the UK and other European countries were delayed by problems with signing up suitable wireless network operators.

Kindle 2



A second generation Kindle

On February 10, 2009, Amazon announced the Kindle 2. It became available for purchase on February 23, 2009. The Kindle 2 features a text-to-speech option to read the text

aloud, and 2 GB of internal memory of which 1.4 GB is user-accessible. By Amazon's estimates the Kindle 2 can hold about 1500 non-illustrated books. Unlike the Kindle First Generation, Kindle 2 does not have a slot for SD memory cards. It was slimmer than the original Kindle.

To promote the new Kindle, author Stephen King made *UR*, his then-new novella, available exclusively through the Kindle Store. On October 22, 2009, Amazon stopped selling the original Kindle 2 in favor of the international version it had introduced earlier in the month.

According to an early review by iFixIt, the Kindle 2 features a Freescale 532 MHz, ARM-11 90 nm processor, 32 MB main memory, 2 GB moviNAND flash storage and a 3.7 V 1530 mAh lithium polymer battery.

On November 24, 2009, Amazon released a firmware update for the Kindle 2 that it said increases battery life by 85% and introduces native PDF support.

On July 8, 2009, Amazon reduced price of the Kindle 2 from the original \$359 to \$299. On October 7, 2009, Amazon further reduced the price of the Kindle 2 to \$259. The Kindle 2 was criticized for its high original retail price, compared to the \$185.49 manufacturing cost estimated by iSuppli.

International version

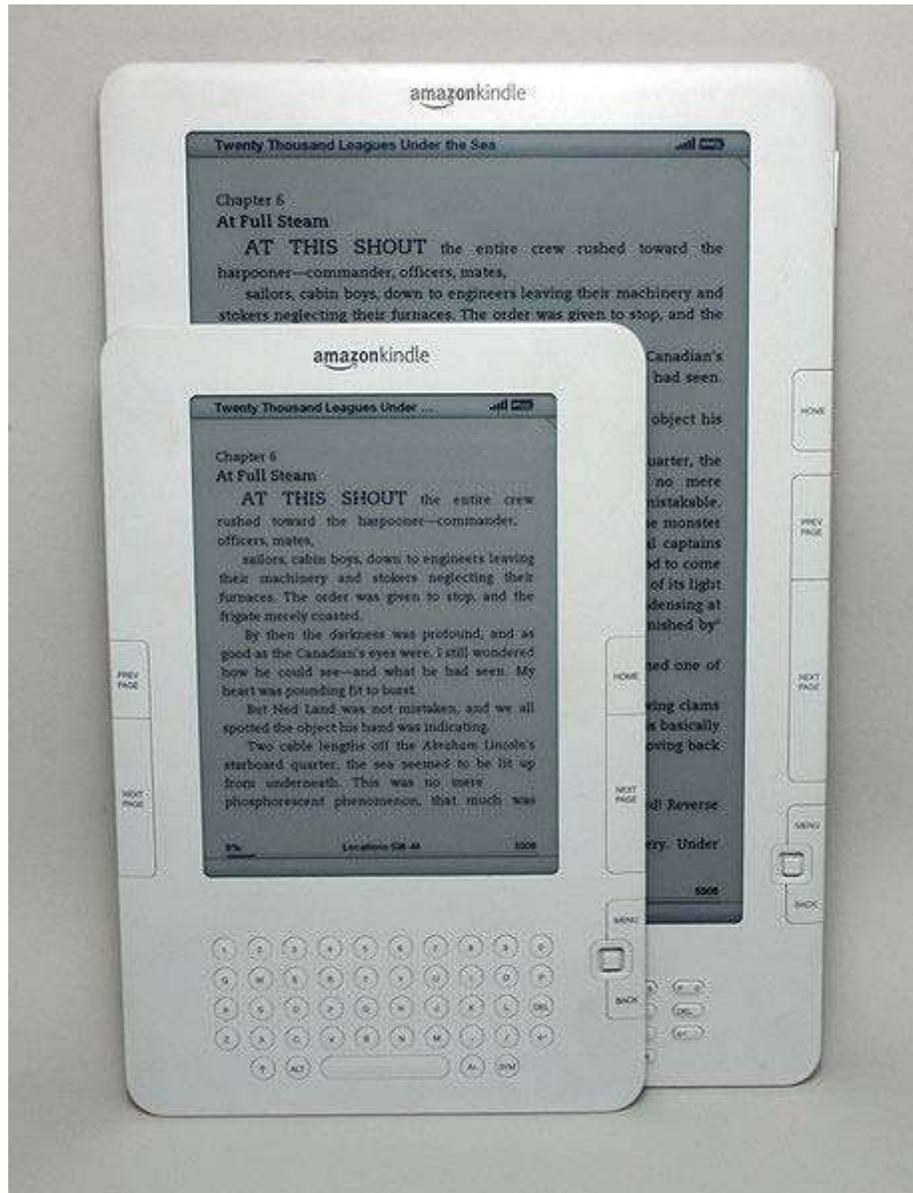
On October 7, 2009, Amazon announced an international version of the Kindle 2 that works in over 100 countries. It became available October 19, 2009. The international Kindle 2 is physically very similar to the U.S.-only model, although it uses a different mobile network standard.

The original Kindle 2 uses CDMA2000, for use on the Sprint network. The international version uses standard GSM and 3G GSM, enabling it to be used on AT&T's U.S. mobile network and internationally in 100 other countries.

Kindle 2 International Version is believed to have a noticeably higher contrast screen, although Amazon does not advertise this.

On October 22, Amazon lowered the price on the international version from \$279 to \$259 and discontinued the U.S.-only model. On June 21, 2010, hours after Barnes & Noble lowered the price of its Nook, Amazon lowered the price of the Kindle 2 to \$189, undercutting the Nook by \$10.

Kindle DX



The larger Kindle DX with a Kindle 2 for size comparison

Amazon announced the Kindle DX on May 6, 2009. This device has a larger screen than the standard Kindle and supports simple PDF files. It was also the thinnest Kindle to date and offers an accelerometer, which enables the user to seamlessly rotate pages between landscape and portrait orientations when the Kindle DX is turned on its side. It is marketed as more suitable for displaying newspaper and textbook content.

International version

Since January 19, 2010, the Kindle DX International has shipped in 100 countries. The Kindle DX comes with a 24.6 cm (9.7 inch) E Ink screen instead of the 15.2 cm (6 inch) normal Kindle screen.

Kindle DX Graphite

On July 1, 2010, Amazon released a new revision of the Kindle DX (3rd Generation Kindle DX). As well as dropping the price from \$489 to \$379, the new Kindle DX has an e-ink screen with 50% better contrast ratio and comes only in a "graphite" color. It is speculated the color change is to improve contrast ratio perception even further, as some users found the previous white casing highlighted the fact that the e-ink background is gray and not white.

Kindle 3 (WiFi and 3G with WiFi, Latest Generation)

Amazon announced a new generation of the Kindle on July 28, 2010. While Amazon does not officially add numbers to the end of each Kindle denoting its generation, most reviewers, customers and press companies refer to this updated Kindle as the "Kindle 3".



Kindle 3. (Left to right) Volume control, headphone jack, microphone, USB and power switch

The Kindle 3 is available in two versions. One of these, the Kindle Wi-Fi, is initially priced at US\$139 / GB£109, and connects to the Internet exclusively via public or private Wi-Fi networks. The other version, considered a replacement to the Kindle 2, is priced at US\$189 / GB£149 and includes both 3G and Wi-Fi connectivity. The built-in free 3G connectivity uses the same wireless signals that cell phones use, allowing it to have download and purchase availability from any location. The new Kindle with 3G is available in two colors: classic white and graphite. Both models use the new E ink "Pearl" display, which Amazon claims is 50% better in contrast—a claim that is backed up by early user reports. Criticism of the Kindle 3 use of 3G technology appears to include less reception than is available with the Kindle 2 3G network.

The Kindle 3 utilizes a Freescale i.MX353 applications processor, Freescale MC13892 power management chip, Epson EINK controller and Samsung DRAM and Flash. Other hardware changes include a larger 1750 mAh lithium polymer battery, AnyDATA DTP-600W 3G GSM modem and Atheros AR6102G 802.11bg WiFi chip.

The third-generation Kindle is 0.5 inches shorter and 0.5 inches narrower than the Kindle 2. It supports additional fonts and international Unicode characters. An experimental browser based on the popular WebKit rendering engine is included, as well as text-to-speech menu navigation. Internal memory is expanded to 4 GB. The battery can allegedly last for up to one month of reading with the wireless radios turned off.

Pre-orders for the new Kindle began concurrent with the announcement of the device, and Amazon began shipping the devices on August 27, 2010 in the United States and United Kingdom.

With the announcement of the Kindle 3, Amazon also launched an Amazon.co.uk version of the Kindle store. As yet it's unclear whether users who move out of the UK will be able to transfer existing purchases to Amazon.com. However, existing UK users are offered the option of migrating to the UK Kindle store, with no loss to their existing purchases.

On August 25, 2010, Amazon announced that the Kindle 3 was the fastest-selling Kindle ever.

Kindle applications

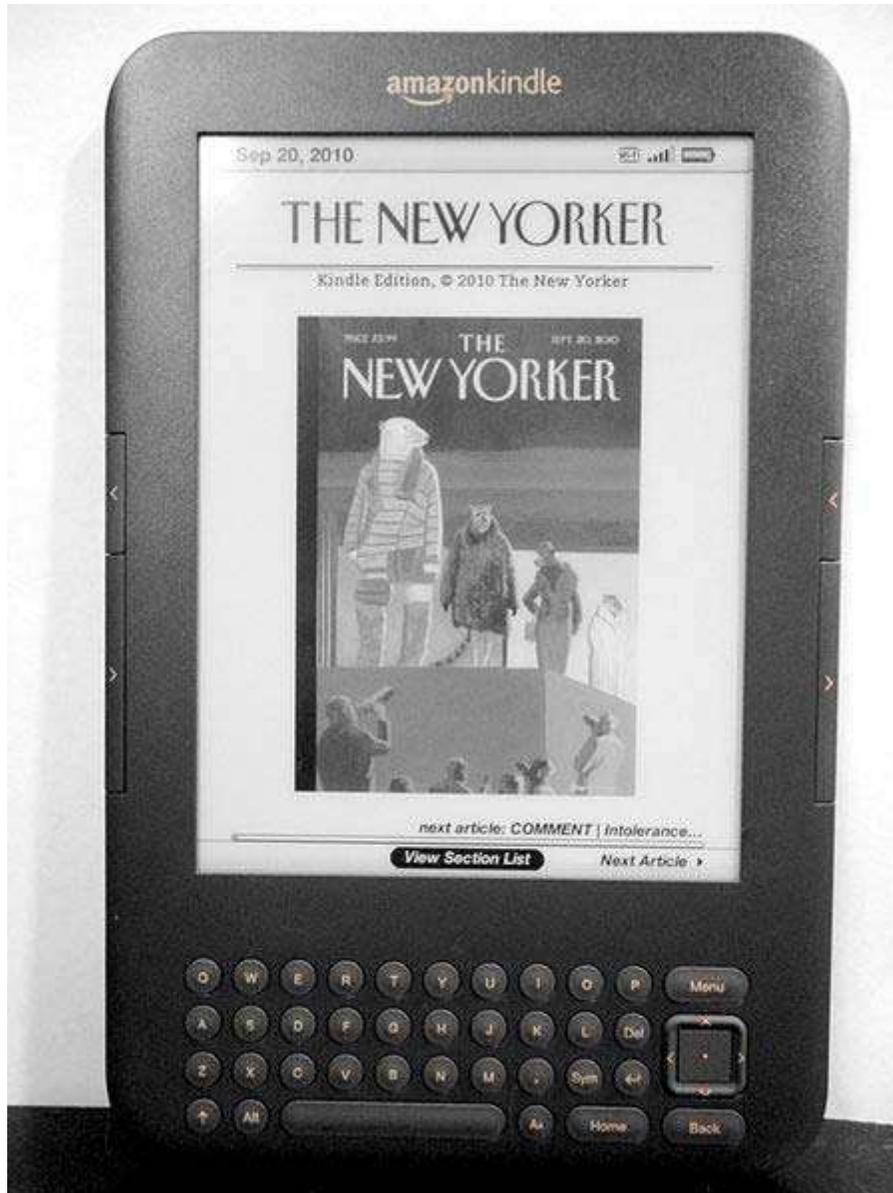
Amazon released a "Kindle for PC" application in late 2009, available as a free download for Windows 7, Vista, and XP. This application allows thousands of books to be read on a personal computer in color, with no Kindle unit required, as e-books can simply be purchased from Amazon's store. Amazon later released a version for the Macintosh, in early 2010. In June 2010, Amazon released a "Kindle for Android" version. With the Android application release, versions for the Apple iPhone, the iPad, PC and Mac computers, and BlackBerry cellphones are also available.

Kindle sales

Specific Kindle sales numbers are not released by the company; however, Jeff Bezos, founder and CEO of Amazon.com, stated in a shareholders' meeting in January 2010 that "millions of people now own Kindles." According to anonymous inside sources, over three million Kindles have been sold as of December 2009, while external estimates, as of Q4-2009, place the number at about 1.5 million. According to James McQuivey of Forrester Research, estimates are ranging around four million, as of mid-2010. They are now being sold for \$139.

Overview

Content from Amazon and some other content providers is primarily encoded in Amazon's proprietary Kindle format (AZW). It is also possible to load content in various formats from a computer by simply transferring it to the Kindle via USB (for free) or by emailing it to a registered email address provided by Amazon (for a fee, unless the transfer to the device is done via Wi-Fi instead of 3G); the email service can convert a number of document formats to Amazon's AZW format and then transmit the result to the associated Kindle over Whispernet.



The New Yorker subscribed on a Kindle 3

Kindle *Terms of Use* forbid transferring Amazon e-books to another user or a different type of device. Users can select reading material using the Kindle itself or through a computer at the Amazon Kindle store and can download content through the Kindle Store, which upon the initial launch of the Kindle had more than 88,000 digital titles available for download. This number continued steadily increasing to more than 275,000 by late 2008, and exceeded 500,000 in the spring of 2010. As of July 24, 2010, there were more than 650,000 books available for download. In late 2007, new releases and *New York Times* best sellers were being offered for approximately US\$11, with first chapters of many books offered as free samples. Many titles, including some classics, are offered free of charge or at a low price, which has been stated to relate to the cost of adapting the book to the Kindle format. Magazines, newspapers and blogs via RSS are provided by

Amazon per a monthly subscription fee or a free trial period. Newspaper subscriptions cost from US\$1.99 to \$27.99 per month; magazines charge between \$1.25 and \$10.99 per month, and blogs charge from \$0.99 to \$1.99 per month. Amazon e-book sales overtook print for one day for the first time on Christmas Day of 2009.

International users of Kindle pay different prices for books depending on their registered country. For U.S. customers traveling abroad, Amazon originally charged a \$1.99 fee to download books over 3G while overseas. That charge was quietly dropped in May 2010. Fees remain for wireless delivery of periodical subscriptions and personal documents.

In addition to the Kindle store, paid content for the Kindle can be purchased from various independent sources such as Fictionwise, Mobipocket and Webscriptions. Public Domain titles are also obtainable for the Kindle via content providers such as Project Gutenberg and World Public Library.

The device is sold with electronic editions of its owner's manual and the *New Oxford American Dictionary*. Users are able to purchase different dictionaries from the Kindle store as specified in the included manual. The Kindle also contains several free experimental features including a basic web browser. Users can also play music from MP3 files in the background in the order they were added to the Kindle. Operating system updates are designed to be received wirelessly and installed automatically during a period in sleep mode in which wireless is turned on.

The Kindle does not preserve the print edition page numbers used for creating and verifying sources, references and citations.

File formats

The original Kindle supported only unprotected Mobipocket books (MOBI, PRC), plain text files (TXT), Topaz format books (TPZ), and Amazon's proprietary DRM-restricted format (AZW). Version 2.3 firmware upgrade for Kindle 2 (U.S. and International) added native Portable Document Format (PDF) support. Earlier versions did not fully support PDF, but Amazon provided "experimental" conversion to the native AZW format, with the caveat that not all PDFs may format correctly. It does not support the EPUB ebook standard. However there is software available (e.g. Calibre) which can convert a non-DRM EPUB file into the unprotected Mobipocket format that the Kindle can read. Amazon offers an email-based service that will convert JPEG, GIF, PNG and BMP graphics to AZW. Amazon will also convert HTML pages and Microsoft Word (DOC) documents through the same email-based mechanism, which will send a Kindle-formatted file to the device directly for \$0.15 per MB or to a personal e-mail account for free. These services can be accessed by sending emails to <kindleusername>@kindle.com and to <kindleusername>@free.kindle.com for Whispernet-delivered and free email-delivered file conversion, respectively, but these are services available just for those who bought a real Kindle device, not available for those who just own the digital Kindle application (iPhone, iPad, etc.). The file that the user wants to be converted needs to be attached to these emails. Users could also convert PDF

and other files to the first-generation Kindle's supported formats using third-party software. The original Kindle supported audio in the form of MP3s and Audible audiobooks (versions 2, 3 and 4), which had to be transferred to the Kindle via USB or on an SD card.

A book may be downloaded from Amazon to a limited number of devices at the same time. The limit ranges from one to six devices, depending on an undisclosed number of licenses set by the book publisher. When the limit is reached, the users have to unregister some devices in the Manage Your Kindle page in order to add new devices.

E-books of unencrypted .MOBI files, .TXT files, or .AZW formats can be transferred to the Kindle over a USB connection and read, but any other e-book formats are not supported. The original Kindle and the Kindle 2 firmware before the 2.3 firmware update cannot read e-books or files in the PDF format. However, PDFs and several other file formats can be converted using a number of downloadable applications, free conversion by email, or a similar method that sends the converted content to the owner's Kindle for a fee.

Amazon owns Mobipocket, and the Kindle AZW file format and DRM scheme are similar to the Mobipocket file format and DRM scheme, yet Kindle is not able to read DRM-protected Mobipocket books without resorting to third-party conversions tools.

Initially, Kindle 1 only supported the ISO 8859-1 (Latin 1) character set for its content; Unicode characters and non-Western characters were not supported. A firmware update in February 2009 added support for additional character sets, including ISO 8859-16.

Kindle 2 added support for Audible Enhanced (AAX) format, but dropped support for Audible versions 2 and 3. Using the experimental web browser, it was possible to download books directly on the Kindle (in MOBI, PRC and TXT formats only). Hyperlinks in a Mobipocket file could be used to download e-books but could not be used to reference books stored in the Kindle's memory. Kindle DX added native support for PDF files.

The original Kindle and Kindle 2 did not allow the user to organize books into folders. There is an option to select whether documents, subscriptions, books, or everything on the device appear on the home page. Another option orders the items on the home page according to title, author, or download date. Books may also be tagged with one or more keywords by inserting the tags into notes added to the book. Users can then search for books by tag. Kindle software version 2.5 (released July 2010) allowed for the organization of books into "Collections" which is roughly correspondent to folders except for the fact that one book may be added to multiple collections.

User-created annotations

Users can bookmark, highlight and look up content. Pages can be *dog-eared* for reference and notes can be added to relevant content. While a book is open on the display, menu

options allow users to search for synonyms and definitions from the built-in dictionary. The device also remembers the last page read for each book. Pages can be saved as a "clipping", or a text file containing the text of the currently displayed page. All clippings are appended to a single file, which can be downloaded over a USB cable.

Kindle Development Kit (KDK)

On January 21, 2010, Amazon announced the forthcoming release of their Kindle Development Kit. Their aim is to allow developers to build 'active content' for the Kindle, and a beta version was announced with a February 2010 release date. A number of companies have already experimented with delivering active content through the Kindle's bundled browser, and the KDK promises 'sample code, documentation and the Kindle Simulator' together with a new revenue sharing model for developers.

The KDK is based on the Java Programming Language, specifically, the JSR 1.1.2 Personal Basis flavor of packaged Java APIs.

Business model

Digital Text Platform

Concurrently with the Kindle device, Amazon launched the Digital Text Platform, a system for authors to self-publish directly to the Kindle. In open beta testing as of late 2007, the platform has been promoted to established authors by e-mail and by advertisements at Amazon.com. Authors can upload documents in several formats for delivery via Whispernet and charge between \$0.99 and \$200.00 per download. The authors receive 35% of revenues based on their list price, regardless of discounts by Amazon.

In a December 5, 2009 interview with *The New York Times*, CEO Jeff Bezos revealed that Amazon.com keeps 65% of the revenue from all ebook sales for the Kindle. The remaining 35% is split between the book author and publisher. After numerous commentators observed that Apple's popular App Store offers 70% of royalties to the publisher, Amazon began a program that offers 70% royalties to Kindle publishers who agree to certain conditions.

Other criticisms involve the business model behind Amazon's implementation and distribution of e-books. Amazon introduced a software application allowing Kindle books to be read on an iPhone or iPod Touch. Amazon soon followed with an application called "Kindle for PCs" that can be run on a Windows PC. Due to the book publisher's DRM policies, there is no right of first sale with the e-books. Amazon states they are licensed, not purchased; so unlike paper books, buyers do not actually own their e-books.

Remote content removal

On July 17, 2009, Amazon.com withdrew certain Kindle titles, *Animal Farm* and *Nineteen Eighty-Four* by George Orwell, from sale, refunded the cost to those who had purchased them, and remotely deleted these titles from purchasers' devices after discovering that the publisher lacked rights to publish the titles in question. Notes and annotations for the books made by users on their devices were left in a separate file, but "rendered useless" without the content they were directly linked to. The move prompted outcry and comparisons to *Nineteen Eighty-Four* itself. In the novel, books, magazines and newspapers in public archives that contradict the ruling party are edited or destroyed, long after being published; the removed materials go "down the Memory Hole", nickname for an incinerator chute. Customers and commentators noted the resemblance to the censorship in the novel, and described Amazon's action in Orwellian terms. Some critics also argued that the deletion violated the Kindle's Terms of Service, which states in part:

"Upon your payment of the applicable fees set by Amazon, Amazon grants you the non-exclusive right to keep a permanent copy of the applicable Digital Content and to view, use and display such Digital Content an unlimited number of times, solely on the Device or as authorized by Amazon as part of the Service and solely for your personal, non-commercial use."

Amazon spokesman Drew Herdener stated that the company is "... changing our systems so that in the future we will not remove books from customers' devices in these circumstances." On July 23, 2009, Amazon CEO Jeff Bezos posted an apology about the company's handling of the matter on Amazon's official Kindle forum. Bezos said the action was "stupid", and that the executives at Amazon "deserve the criticism received."

On July 30, 2009, Justin Gawronski, a Michigan high-school senior, and Antoine Bruguier, a California engineer, filed suit against Amazon in the United States District Court for the Western District of Washington. Gawronski argued that Amazon had violated their terms of service by remotely deleting the copy of *Nineteen Eighty-Four* he had purchased, in the process preventing him from accessing annotations he had written. Bruguier also had his copy deleted without his consent, and found Amazon "deceit[ful]" in an email exchange. The complaint, which requested class-action status, asked for both monetary and injunctive relief. The case was settled on September 25, 2009, with Amazon agreeing to pay \$150,000 divided between the two plaintiffs, on the understanding that the law firm representing them, Kamber Edelson LLC, "...will donate its portion of that fee to a charitable organization...". The settlement also saw Amazon guaranteeing wider rights to Kindle owners over their eBooks:

For copies of Works purchased pursuant to TOS granting "the non-exclusive right to keep a permanent copy" of each purchased Work and to "view, use and display [such Works] an unlimited number of times, solely on the [Devices]. . . and solely for [the purchasers'] personal, non-commercial use", Amazon will not remotely delete or modify such Works from Devices purchased and being used in the United States unless (a) the user consents

to such deletion or modification; (b) the user requests a refund for the Work or otherwise fails to pay for the Work (e.g., if a credit or debit card issuer declines to remit payment); (c) a judicial or regulatory order requires such deletion or modification; or (d) deletion or modification is reasonably necessary to protect the consumer or the operation of a Device or network through which the Device communicates (e.g., to remove harmful code embedded within a copy of a Work downloaded to a Device).

On September 4, 2009, Amazon offered affected users a restoration of the deleted ebooks, an Amazon gift certificate, or a check for the amount of \$30.

Kindle accessories

Amazon also offers a number of peripherals and accessories for all the Kindle models. Kindle covers, earphones, reading lights, and stands being the most popular. Kindle also offers snap-on lights, storage sleeves, waterproof cases, car chargers, microphones, bags, and screen protectors.

The Kindle cover is very popular, and comes for all models of the Kindle reader. Made of leather, many of the covers come in different colors, and are splash-proof. Some of these Kindle covers include a pocket to store a reading light. Other jackets have built-in lights.

Chapter 5

PocketBook Reader

PocketBook is a series of electronic devices for reading electronic books, RSS feeds, and listening of audio books and other MP3 audio files. It is manufactured by the PocketBook International, a Hong Kong-based multinational company with Headquarters in Kyiv, Ukraine. Major manufacturing facilities are located in Taiwan. USA headquarters are located in Overland Park, Kansas (a Kansas City suburb).

Description

PocketBook is a family of devices for electronic book reading, where most of the devices use the E-Ink Corporation electronic ink technology. Vizplex display provides the convenience of reading (no flicker and no letters shape changes, regardless of lighting conditions and angle of view, the ability to read in direct sunlight) and ultra low power consumption. 16 shades of gray allow you to view black and white images. The battery lasts without being recharged for many thousands of page turnings. As the device does not consume energy when displaying an already opened page, the energy consumption is usually calculated by the number of page turnings.

When connected to a PC, these devices are recognized as a USB-drive, which allows easy transfers of electronic books to the device from popular operating systems.

The PocketBook E-Ink models compete with other e-paper devices: Sony Reader, Amazon Kindle II, etc.

Supported file formats

- Book formats:
 - CHM, DOC, DJVU, EPUB, FB2, FB2.ZIP, HTML, PDF, PRC (Mobipocket), RTF, TCR, TXT
 - DRM file formats: EPUB, PDF
- Audio formats:
 - MP3 (No sound in 360 model)
 - AAC, WMA (701 model only)
- Image formats:
 - BMP, JPEG, PNG, TIFF

- Video formats:
 - 3GP, AVI, MP4 (701 model only)
- Internet formats:
 - Web browser (701, 602/603, 902/903 models only)

TFT LCD models

PocketBook IQ 701

- Display: 7" TFT LCD, 600x800, 143 dpi, Color
- Touch Panel: Yes
- Operating System: Android 2.0
- CPU: Samsung, 800MHz
- Internal Memory: RAM 256 Mb, Flash 2 Gb
- External Memory Slot: SD (up to 2 Gb) or SDHC (up to 32 Gb)
- Wireless Connectivity: 802.11 b/g/n
- Wired Connectivity: Mini-USB 2.0
- Audio Output: 3.5mm stereo TRS connector, 2 speakers
- Accelerometer: Yes
- Battery: Li-Polymer, 1600 mAh
- Size: 144 x 192 x 14 mm
- Weight: 516 grams (1.14 lb)
- Color: Red, White, or Blue
- Wall Charger: Yes; OUTPUT 12V, 1.5A
- Approximate Price: **USD \$139**

E-Ink models

PocketBook Pro 903

Same as 902, **plus:**

- Wireless Connectivity: 3G Wireless
- Touch Panel: Yes
- Weight: 581 grams (1.28 lb)
- Color: Dark Silver only
- Approximate Price: Not Shipping Yet?

PocketBook Pro 902

- Display: 9.7" E-Ink, 825x1200, 166 dpi, B&W
- Touch Panel: None
- Operating System: Linux 2.6.28
- CPU: Samsung, 533 MHz
- Internal Memory: RAM 256 Mb, Flash 2 Gb

- External Memory Slot: microSD (up to 2 Gb) or microSDHC (up to 32 Gb)
- Wireless Connectivity: 802.11 b/g, Bluetooth
- Wired Connectivity: Micro-USB
- Audio Output: 3.5mm stereo TRS connector, 2 speakers
- Accelerometer: Yes
- Battery: Li-Polymer, 1530 mAh
- Size: 192 x 263 x 11.5 mm
- Weight: 530 grams (1.2 lb)
- Color: Dark Grey or White
- Wall Charger: Option
- Approximate Price: **USD \$299**

PocketBook Pro 603

Same as 602, **plus:**

- Wireless Connectivity: 3G Wireless
- Touch Panel: Yes
- Weight: 280 grams (0.62 lb)
- Color: Dark Silver only
- Approximate Price: Not Shipping Yet?

PocketBook Pro 602

- Display: 6" E-Ink, 600x800, 166 dpi, B&W
- Touch Panel: None
- Operating System: Linux 2.6.28
- CPU: Samsung, 533 MHz
- Internal Memory: RAM 256 Mb, Flash 2 Gb
- External Memory Slot: microSD (up to 2 Gb) or microSDHC (up to 32 Gb)
- Wireless Connectivity: 802.11 b/g, Bluetooth
- Wired Connectivity: Micro-USB
- Audio Output: 3.5mm stereo TRS connector, 2 speakers
- Accelerometer: Yes
- Battery: Li-Polymer, 1530 mAh
- Size: 138 x 182 x 10.6 mm
- Weight: 250 grams (0.55 lb)
- Color: Dark Grey or White
- Wall Charger: Option
- Approximate Price: **USD \$169**

PocketBook 360



PocketBook 360

- Display: 5" E-Ink, 600x800, 200 dpi, B&W
- Touch Panel: None
- Operating System: Linux
- CPU: Samsung, S3C2440AL-40, ARM9 (ARM920T), 400MHz
- Internal Memory: RAM 64 Mb, Flash 512 Mb, User-accessible 466 Mb
- External Memory Slot: microSD (up to 2 Gb) or microSDHC (up to 32 Gb)
- Wireless Connectivity: None
- Wired Connectivity: Mini-USB
- Audio Output: **None**, thus audio files are not supported

- Accelerometer: Yes
- Battery: Li-Polymer, 1000 mAh
- Size: 118 x 140 x 10 mm, or 118 x 140 x 12 mm with cover
- Weight: 150 grams (0.33 lb)
- Color: Black or Ivory
- Wall Charger: Yes
- Approximate Price: **USD \$159**

PocketBook 302

- Display: 6" E-Ink, 600x800, 166 dpi, B&W
- Touch Panel: Yes
- Operating System: Linux
- CPU: Samsung, S3C2440AL-40, ARM9 (ARM920T), 400MHz
- Internal Memory: RAM 64 Mb, Flash 1 Gb, User-accessible 466 Mb
- External Memory Slot: microSD (up to 2 Gb) or microSDHC (up to 32 Gb)
- Wireless Connectivity: WLAN, Bluetooth
- Wired Connectivity: Mini-USB
- Audio Output: 3.5mm stereo TRS connector
- Accelerometer: Yes
- Battery: Li-Polymer, 2050 mAh
- Size: 131 x 190 x 13 mm
- Weight: 280 grams (0.62 lb)
- Color: Black or White
- Wall Charger: Yes
- Approximate Price: **USD \$199**

PocketBook 301+



PocketBook 301

- Display: 6" E-Ink, 600x800, 166 dpi, B&W
- Touch Panel: None
- Operating System: Linux
- CPU: Samsung, S3C2440AL-40, ARM9 (ARM920T), 400MHz
- Internal Memory: RAM 64 Mb, Flash 512 Mb
- External Memory Slot: SD (up to 2 Gb) or SDHC (up to 32 Gb)
- Wireless Connectivity: None
- Wired Connectivity: Mini-USB
- Audio Output: 2.5mm stereo TRS connector
- Accelerometer: None
- Battery: Li-Polymer, 1000 mAh
- Size: 118 x 188 x 8.5 mm
- Weight: 174 grams (0.38 lb)
- Color: Black, Grey, Pink, or Light-Green
- Wall Charger: Yes; INPUT 100-240V, 50/60Hz, 0.2A; OUTPUT 5V, 0.5-1A
- Approximate Price: **USD \$169**

PocketBook 301

- Older Model

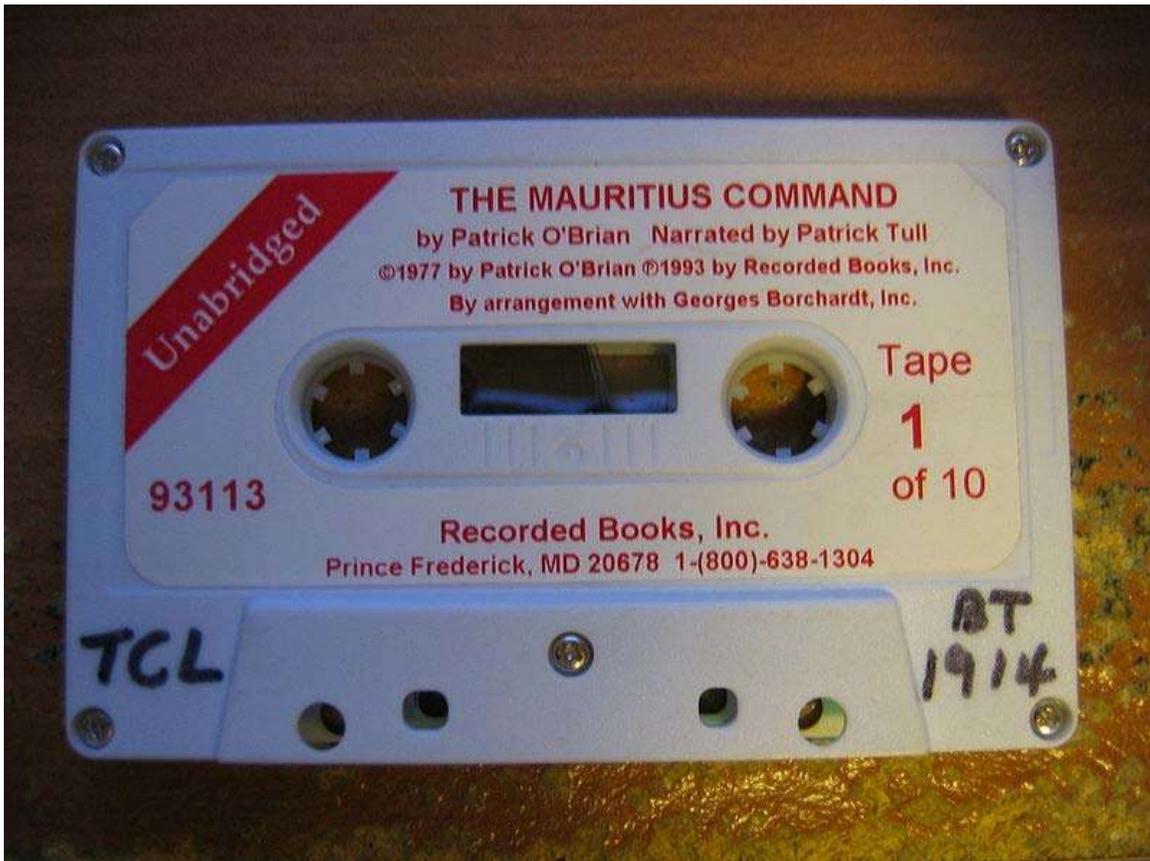
Languages

2010-year models support nearly 20 languages for menu and virtually any language for e-reading including Asian languages. Translation between English and 24 other European languages is available as well.

WWT

Chapter 6

Introduction to Audiobook



An audio cassette recording

An **audiobook** is a recording that is primarily spoken word. It is often based on a recording of commercial printed material. It is not necessarily an exact audio version of a book.

Spoken audio was available in school and public libraries and to a lesser extent in music shops. It was not until the 1980s that there began a concerted effort to attract book retailers. As publishers entered the field of spoken-word publishing, the transition to book

retailers carrying audiobooks became commonplace on bookshelves rather than in separate displays.

Formats

Audiobooks are distributed on CDs, cassette tapes, downloadable digital formats (e.g., MP3 (.mp3), Windows Media Audio (.wma), and Advanced Audio Coding (.aac)). A benefit to the CD format is that CDs can be ripped and then put onto an MP3 player; a 14 CD "book" can be put onto a tiny & cheap (\$10USD) MP3 player.

In 2005 cassette-tape sales were 16% of the audiobook market, with CD sales accounting for 74% of the market and downloadable audio books accounting for approximately 9%. In the United States, a sales survey (performed by the Audio Publishers' Association in the summer of 2006 for the year 2005) estimated the industry to be worth 871 million US dollars. Current industry estimates are around two billion US dollars at retail value per year. In recent years, the Internet has introduced another powerful means of delivery for audiobooks and many titles are now available on-line, as downloads and as audio streams.

Sometimes audiobook format is available simultaneously with book publication. There are 50,000 titles on cassette, CD or digital format.

Unabridged audiobooks are word for word readings of a book, while abridged audiobooks have text removed by the abridger. Abridgements may be wanted to reduce the cost or for other reasons, but are also criticized for being incomplete versions of the original work. The increasing use of digital formats for audiobook recording has led to much less abridgment of texts, the most common reason cited for abridgment was the cost of production and with digital recordings being stored as data rather than on discs or tapes, the differential in cost of producing the full text was minimised. Audiobooks may come as fully dramatized versions of the printed book, sometimes calling upon a complete cast, music, and sound effects, though many consumers have indicated a preference for less music and fewer multiple voices and sound effects. Each spring, the Audie Awards are given to the top nominees for performance and production in several genre categories. Relatedly, a dramatized audio adaptation of a book is one form of an audio drama.

Occasionally there are radio programs serializing books, sometimes read by the author or sometimes by an actor, with most of them on the BBC.

History

United States

In 1931, Congress established the talking-book program, which was intended to help blind adults who couldn't read print. This program was called "Books for the Adult Blind

Project." The American Foundation for the Blind developed the first talking books in 1932. One year later the first reproduction machine began the process of mass publishing. In 1933 anthropologist J.P. Harrington drove the length of North America to record oral histories of Native American tribes on aluminum discs using a car battery-powered turntable. Audiobooks preserve the oral tradition of storytelling that J.P. Harrington pursued many years ago. By 1935, after Congress approved free mailings of audio books to blind citizens, the Books for the Adult Blind Project was in full operation. In 1992 the National Library Service for the Blind and Physically Challenged (NLS) network circulated millions of recorded books to more than 700,000 Physically challenged listeners. All NLS recordings were created by professionals.

Though spoken recordings were popular in 33-1/3 vinyl record format for schools and libraries into the early 1970s, the beginning of the trade acceptance of this medium can be traced to the introduction of the audio cassette and, most importantly, to the prevalence of these cassette players as standard equipment (rather than as options which older drivers did not choose) in imported (Japanese) automobiles, which became very popular during the 1979 energy crisis. Thereafter, consumers and authors slowly accepted the medium. Into the early 1980s there were still many authors who refused to have their books created as audiobooks, so a good many of the audiobooks were original productions not based upon printed books.

With the development of portable cassette recorders, audiotapes had become very popular and by the late 1960s libraries became a source of free audiobooks, primarily on vinyl records but also on cassettes. Instructional and educational recordings came first, followed by self-help tapes and then by literature. In 1975, Olympic gold medalist, Duvall Hecht founded Books on Tape, Inc. as a direct to consumer mail order rental service for unabridged audiobooks and expanded their services selling their products to libraries and audiobooks gaining popularity with commuters and travelers. By the middle of 1980s the audio publishing business grew to several billion dollars a year in retail value.

Caedmon was the first to work with integrated production teams and professional actors, while Nightingale Conant featured business and self-help authors reading their own works first on vinyl records and then on cassettes.

The Audio Publishers Association was established in 1986 by six competitive companies who joined together to promote the consumer awareness of spoken word audio. In 1996 the Audio Publishers Association established the Audie Awards for audio books, which is equivalent to the Oscar for the talking books industry. The nominees are announced each year in January. The winners are announced at a gala banquet in the spring, usually in conjunction with BookExpo America.

While music fans rapidly accepted CDs, audiobook listeners were slower. One reason may have been that a cassette tape by nature retains its position when the player is turned off, but many early CD players did not retain the playing position of CDs when turned

off. Also, it was not until cassette players were replaced by CD players in most automobiles that this format eventually took hold.

With the advent of the Internet, broadband technologies, new compressed audio formats and portable media players, the popularity of audiobooks has increased significantly. This growth was reflected with the advent of audiobook download subscription services.

Use, distribution and popularity

Recent technology has encouraged the proliferation of free audiobooks that take works from the public domain and enlist volunteers to read them. Audiobooks also can be created with text to speech computer software, although the quality of synthesised speech may suffer by comparison to recordings by a human voice. On the other hand, computer-voiced reading enables the proliferation of more works faster through automation, than if read by humans.

Audiobooks in the private domain are also distributed online by for-profit companies. It is believed that most major audiobook publishers insist that their works, when sold as downloads, be protected by Digital Rights Management (DRM), but this is not the case. Companies such as Apple Inc. have licensed their proprietary FairPlay DRM system (for DRM protection of iPod files in the .aa file format) to only one company. . Because of the major publishers' insistence on DRM, this has effectively created a monopoly in the sale of the works of major publishers to iPod users, who make up the majority of the portable audio market. However, with the release of new sites such as We Read 4 You, which releases all of its content as DRM-free .mp3 files, there may be a slow conversion by other companies, such as Apple and Amazon to a DRM-free format.

Audiobooks on cassette or CD are typically more expensive than hardbacks because of the added expense of recording and the lack of the economy of scale in high "print" runs that are available in the publishing of printed books. Preloaded digital formats are similar in price to their CD counterparts. The audio content is preloaded on a small and simple player, which removes the need for a separate piece of technology such as a CD player or an MP3 player.

Downloadable audiobooks tend to cost slightly less than hardbacks but more than their paperback equivalents. For this reason, market penetration of audiobooks is substantially lower than for their printed counterparts despite the high market penetration of the hardware (MP3 and WMA players) and despite the massive market penetration achieved by audio music products. Given the elasticity of demand for audiobooks and the availability of cheaper alternatives, slow and steady growth in sales seems more likely than a mass market explosion.

However, economics are on the side of downloadable audiobooks in the long run. They do not carry mass production costs, do not require storage of a large inventory, do not require physical packaging or transportation and do not face the problem of returns that add to the cost of printed books. Received wisdom of market forces suggests that

significant price reductions to customers, while cutting into per unit profit margins, will be offset by increased volumes of sales. This will increase absolute profits to the industry while bringing audiobooks to a wider public.

One of the factors holding back price competition is the fear that low-price audiobooks might simply take business away from more traditional forms of publishing. This is especially significant in the case of publishers who have interests in both print and audiobook publishing. However, most major book publishers now actively participate in audiobook publishing and see it as a complement to their publishing operations.

Resellers of audiobooks that acquire much of their content from major publishers, must price their content at such a level as to take account of their cost of goods as well as operating costs. On the other hand, audiobook sellers that sell their own content or publish lesser known authors have lower operating costs and can therefore sell at lower prices using a "lower-margin-higher-sales" business model. However, they still have to meet the costs of writer's royalties, performers fees and production facility costs. The shift from CDs and cassettes to downloadable audiobooks, whilst doing nothing to reduce initial recording and editing costs, creates further downward pressure on price, by removing some of the other costs, such as production, packaging and physical distribution.

Audiobooks have been used to teach children to read and to increase reading comprehension. They are also useful for the blind. The National Library of Congress in the U.S. and the CNIB Library in Canada provide free audiobook library services to the visually impaired; requested books are mailed out (at no cost) to clients.

About forty percent of all audiobook consumption occurs through public libraries, with the remainder served primarily through retail book stores. Library download programs are currently experiencing rapid growth (more than 5,000 public libraries offer free downloadable audio books). Libraries are also popular places to check out audio books in the CD format. According to the National Endowment for the Arts' recent study, "Reading at Risk", audio book listening is one of very few "types" of reading that is increasing general literacy.

Listening practices

Audio books are considered a valuable learning tool because of their format. Unlike traditional books or a video program, one can learn from an audiobook while doing other tasks, although it should be noted that this can detract from the primary task, assuming the learning is not the main activity. Such multitasking is feasible when doing mechanical tasks that do not require much thought and have only little or no chance of an emergency arising. Such tasks include doing the laundry and exercising indoors, among others. The most popular general use of audiobooks by adults is when driving an automobile or traveling with public transport, as an alternative to radio. Many people listen as well just to relax or as they drift off into sleep.

Common practices include:

- **Replaying:** Depending upon one's degree of attention and interest, it is often necessary to listen to segments of an audio book more than once to allow the material to be understood and retained satisfactorily. Replaying may be done immediately or after extended periods of time.
- **Learning:** People may listen to an audio book (usually an unabridged one) while following along in an actual book. This helps them to learn words that they may not learn correctly if they were only to read the book. This can also be a very effective way to learn a new language.

Audiobook charities in the UK

Listening Books is an audiobook charity in the UK providing an internet streaming and postal service to anyone who has a disability or illness which makes it difficult to hold a book, turn its pages, or read in the usual way. They have audiobooks for both leisure and learning and a library of over 4,000 titles which are recorded in their own digital studios or commercially sourced.

The Royal National Institute of Blind People (RNIB) is a UK charity which offers a Talking Books library service. The audio books are provided in DAISY format and delivered to the reader's house by post. There are over 18,000 audio books available to borrow, paid for by annual subscription. RNIB subsidises the Talking Books service by around £4 million a year.

Spoken word album

A **spoken word album** was a record album that did not consist mainly of music or songs, but of spoken material. It could be said to be the ancestor of today's audiobook format. Spoken word albums ranged from such items as recordings of actual political speeches and/or dramatic readings of historical documents, to dialogue from the soundtrack of a film, to condensed dramatized versions of literary classics, to complete performances of plays by Shakespeare or other great authors, to stories for children, or to standup comedy routines recorded live in nightclubs.

Early beginnings

Spoken word albums have been made since the early days of recording; examples include the popular Ronald Colman 1941 version of Charles Dickens's *A Christmas Carol* on American Decca Records. However, a true milestone was reached when Columbia Masterworks, which had previously released an album of excerpts from Shakespeare's *Richard II* with Maurice Evans, recorded most of Margaret Webster's famed (and never filmed) 1943 Broadway production of *Othello*, starring Paul Robeson, José Ferrer, and Uta Hagen, on an eighteen-record 78-RPM set that was later transferred to LP. It was the longest spoken word album made up to that time. The album gave millions of listeners

who otherwise were unable to attend a theatrical performance a chance to hear Robeson as Othello and Ferrer as Iago. Sales of the album, however, were affected after Robeson was blacklisted by the U.S. government in 1950 for his alleged Communist sympathies.

LP influence and educational value

After the advent of LPs, spoken word albums became much more common. The Ronald Colman *A Christmas Carol* was transferred to LP, as were many other 78-RPM spoken word albums made by Decca, such as *Moby Dick*, with Charles Laughton as Captain Ahab. An uncredited actor provided the voice of Ishmael, the narrator. Disneyland Records issued many spoken word albums for children, including narrated adaptations of the Disney classic films in their "Disneyland Storyteller" series. Notable Broadway productions, such as the 1950 *Don Juan in Hell* (with Charles Boyer, Charles Laughton, Cedric Hardwicke and Agnes Moorehead), the 1953 dramatized reading of the poem *John Brown's Body* (with Tyrone Power, Judith Anderson and Raymond Massey), the original 1962 Broadway version of *Who's Afraid of Virginia Woolf?* (with Uta Hagen, Arthur Hill, George Grizzard and Melinda Dillon), and Richard Burton's Broadway production of *Hamlet* (1964), were all recorded complete with their original casts by Columbia Masterworks. José Ferrer was heard in an album of excerpts from *Cyrano de Bergerac*, with members of the cast of the 1946 Broadway revival in which Ferrer first played Cyrano, on Capitol Records. Caedmon Records recorded the complete plays of Shakespeare, as well as recordings of other plays such as *Death of a Salesman* with original stars Lee J. Cobb and Mildred Dunnock, *The Glass Menagerie* with Jessica Tandy, and a nearly complete *Cyrano de Bergerac* with Ralph Richardson, who had triumphed in the role in London in 1946. Many of these recorded dramas were played in high school literature classes to enable students to hear the play and follow along in their textbooks at the same time.

After the assassination of President John F. Kennedy in 1963, memorial collections of his speeches began to appear on LP. Most of the soundtrack of the commemorative 1966 documentary *John F. Kennedy: Years of Lightning, Day of Drums*, narrated by Gregory Peck, was released on a Capitol Records LP.

Dialogue excerpts were also released of the film soundtracks of Franco Zeffirelli's *The Taming of the Shrew* (1967) (on RCA Victor), and *Romeo and Juliet* (1968) (on Capitol). RCA Victor also released a virtually complete 2-LP album of the film soundtrack of *A Man for All Seasons*. RCA Victor was also responsible for complete 4-LP album sets of the Laurence Olivier *Othello* and Zeffirelli's National Theatre of Great Britain production of *Much Ado About Nothing*, starring Maggie Smith, Robert Stephens, and Albert Finney. Comedy routines by such celebrities as Bill Cosby became extremely popular on Warner Bros. Records. Emlyn Williams recorded an edition of his one-man performance as Charles Dickens (for Argo Records), and Hal Holbrook recorded excerpts from his one-man *Mark Twain Tonight!* for Columbia Masterworks.

Decline

With the advent of videocassettes and compact discs, however, original cast albums of non-musical plays, as well as spoken word albums of film soundtracks, went into a serious decline from which they have never completely recovered. CDs usually place more emphasis on music than on the spoken word, and there was little interest in only listening to a play or dialogue excerpts from a film when one could now buy plays and films on video and watch them at home whenever they wished. While the Cosby albums have resurfaced on CD, most of the other albums mentioned above have not. (Some of the Caedmon albums have been released on CD by Harper Audio, a division of Harper Collins, which now owns Caedmon.) The 1968 album of *Romeo and Juliet* excerpts has also appeared on CD, and Pearl has issued the Robeson *Othello* in that medium, but the CD edition of the *Othello* has, unfortunately, attracted little attention in comparison to the history-making vinyl record release of the 1940s, and now that *Cyrano de Bergerac*, *A Man for All Seasons*, the Olivier *Othello*, the Zeffirelli versions of *Romeo and Juliet* and *The Taming of the Shrew*, the television version of *Mark Twain Tonight*, and *Richard Burton's Hamlet* are all available on DVD, this has become for most a more preferred way to experience these productions.

Although Naxos Records is a major producer of audiobooks, many famous spoken word recordings of the past, such as Columbia Masterworks' *John Brown's Body* and *Don Juan in Hell* have yet to be released on CD, although *Don Juan in Hell* has become available as an mp3 download. Whether or not it will appear in CD form is still unknown. Also online (but not yet on CD) is Capitol Records' *The Story Teller: A Session with Charles Laughton*, a Grammy-winning one-man stage performance by the actor, featuring dramatic readings from the Bible, Shakespeare, George Bernard Shaw and Jack Kerouac, as well as autobiographical reminiscences.

Today's spoken word albums

There have been some spoken word albums over the past fifteen years or so that have been recorded specifically for compact disc; these have often been combined with classical music. Among them are the Naxos audiobooks, as well as a Chandos Records series of albums which combine the music that William Walton wrote for several Shakespeare productions (including the Olivier film adaptations) with readings from the author performed by such actors as John Gielgud and Christopher Plummer. There is also a Hyperion Records stereophonic re-creation of Ralph Vaughan Williams's 1942 radio play adaptation of *The Pilgrim's Progress*, again with Gielgud.

Today, such websites as BBC, ZBS, L.A. Theatre Works and The Hollywood Theater of the Ear offer full-length recordings on CD of their dramatic productions. These recordings are possibly the closest that modern day discs have come to the spoken word albums of the 1960's.

Chapter 7

Windows Media Audio

Windows Media Audio



Filename extension	.wma
Internet media type	audio/x-ms-wma
Uniform Type Identifier	com.microsoft.windows-?media-wma
Developed by	Microsoft

Windows Media Audio (WMA) is an audio data compression technology developed by Microsoft. The name can be used to refer to its audio file format or its audio codecs. It is a proprietary technology that forms part of the Windows Media framework. WMA consists of four distinct codecs. The original WMA codec, known simply as *WMA*, was conceived as a competitor to the popular MP3 and RealAudio codecs. *WMA Pro*, a newer and more advanced codec, supports multichannel and high resolution audio. A lossless codec, *WMA Lossless*, compresses audio data without loss of audio fidelity (the regular WMA format is not lossless). And *WMA Voice*, targeted at voice content, applies compression using a range of low bit rates.

Development history

The first WMA codec was based on earlier work by Henrique Malvar and his team which was transferred to the Windows Media team at Microsoft. Malvar was a senior researcher and manager of the Signal Processing Group at Microsoft Research, whose team worked on the *MSAudio* project. The first finalized codec was initially referred to as *MSAudio 4.0*. It was later officially released as *Windows Media Audio*, as part of Windows Media Technologies 4.0. Microsoft claimed that WMA could produce files that were half the size of equivalent-quality MP3 files; Microsoft also claimed that WMA delivered "near CD-quality" audio at 64 kbit/s. The former claim however was rejected by some

audiophiles. RealNetworks also challenged Microsoft's claims regarding WMA's superior audio quality compared to RealAudio.

Newer versions of WMA became available: *Windows Media Audio 2* in 1999, *Windows Media Audio 7* in 2000, *Windows Media Audio 8* in 2001, and *Windows Media Audio 9* in 2003. Microsoft first announced its plans to license WMA technology to third-parties in 1999. Although earlier versions of Windows Media Player played WMA files, support for WMA file creation was not added until the seventh version. In 2003, Microsoft released new audio codecs that were not compatible with the original WMA codec. These codecs were *Windows Media Audio 9 Professional*, *Windows Media Audio 9 Lossless*, and *Windows Media Audio 9 Voice*.

Container format

A WMA file is in most circumstances encapsulated, or contained, in the Advanced Systems Format (ASF) container format, featuring a single audio track in one of following codecs: WMA, WMA Pro, WMA Lossless, or WMA Voice. These codecs are technically distinct and mutually incompatible. The ASF container format specifies how metadata about the file is to be encoded, similar to the ID3 tags used by MP3 files. Metadata may include song name, track number, artist name, and also audio normalization values.

This container can optionally support digital rights management (DRM) using a combination of elliptic curve cryptography key exchange, DES block cipher, a custom block cipher, RC4 stream cipher and the SHA-1 hashing function.

Codecs

Windows Media Audio

Windows Media Audio (WMA) is the most common codec of the four WMA codecs. Colloquial usage of the term *WMA*, especially in marketing materials and device specifications, usually refers to this codec only. The first version of the codec released in 1999 is regarded as WMA 1. In the same year, the bit stream syntax, or compression algorithm, was altered in minor ways and became WMA 2. Since then, newer versions of the codec were released, but the decoding process remained the same, ensuring compatibility between codec versions. WMA is a lossy audio codec based on the study of psychoacoustics. Audio signals that are deemed to be imperceptible to the human ear are encoded with reduced resolution during the compression process.

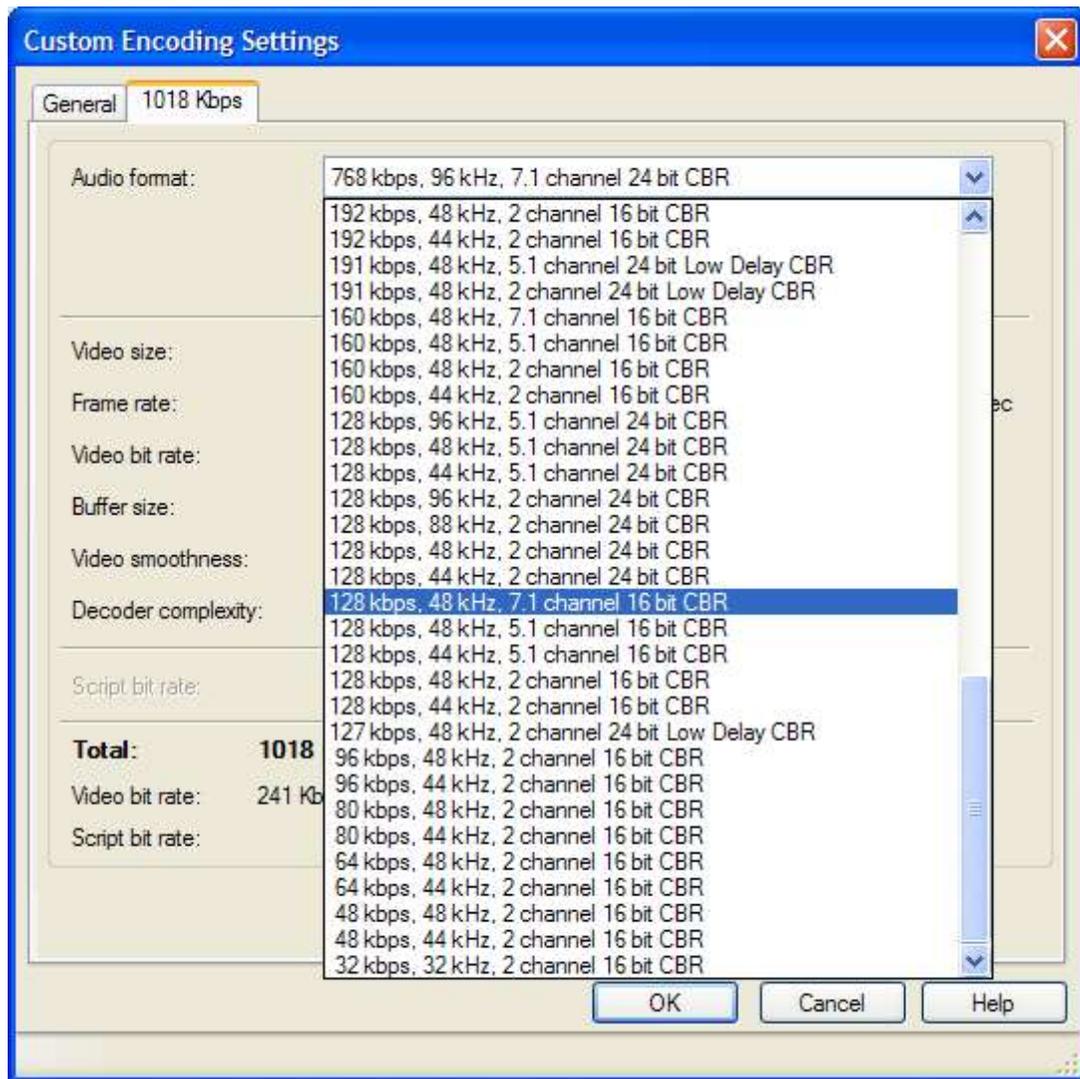
WMA can encode audio signals sampled at up to 48 kHz with up to two discrete channels (stereo). WMA 9 introduced variable bit rate (VBR) and average bit rate (ABR) coding techniques into the MS encoder although both were technically supported by the original format. WMA 9.1 also added support for low-delay audio, which reduces latency for encoding and decoding.

Fundamentally, WMA is a transform coder based on modified discrete cosine transform (MDCT), somewhat similar to AAC, Cook and Vorbis. The bit stream of WMA is composed of superframes, each containing 1 or more frames of 2048 samples. If the bit reservoir is not used, a frame is equal to a superframe. Each frame contains a number of blocks, which are 128, 256, 512, 1024, or 2048 samples long after being transformed into the frequency domain via the MDCT. In the frequency domain, masking for the transformed samples is determined, and then used to requantize the samples. Finally, the floating point samples are decomposed into coefficient and exponent parts and independently huffman coded. Stereo information is typically mid/side coded. At low bit rates, line spectral pairs (typically less than 17 kbit/s) and a form of noise coding (typically less than 33 kbit/s) can also be used to improve quality.

Like AAC and Ogg Vorbis, WMA was intended to address perceived deficiencies in the MP3 standard. Given their common design goals, it's not surprising that the three formats ended up making similar design choices. All three are pure transform codecs. Furthermore the MDCT implementation used in WMA is essentially a superset of those used in Ogg and AAC such that WMA iMDCT and windowing routines can be used to decode AAC and Ogg Vorbis almost unmodified. However, quantization and stereo coding is handled differently in each codec. The primary distinguishing trait of the WMA Standard format is its unique use of 5 different block sizes, compared to MP3, AAC, and Ogg Vorbis which each restrict files to just two sizes. WMA Pro extends this by adding a 6th block size used at 88.1/96 kHz sampling rate.

Certified PlaysForSure devices, as well as a large number of uncertified devices, ranging from portable hand-held music players to set-top DVD players, support the playback of WMA files. Most PlaysForSure-certified online stores distribute content using this codec only. In 2005, Nokia announced its plans to support WMA playback in future Nokia handsets. In the same year, an update was made available for the PlayStation Portable (version 2.60) which allowed WMA files to be played on the device for the first time.

Windows Media Audio Professional



Screenshot of Windows Media Encoder 9 Series, displaying new encoding options for Windows Media Audio 10 Professional.

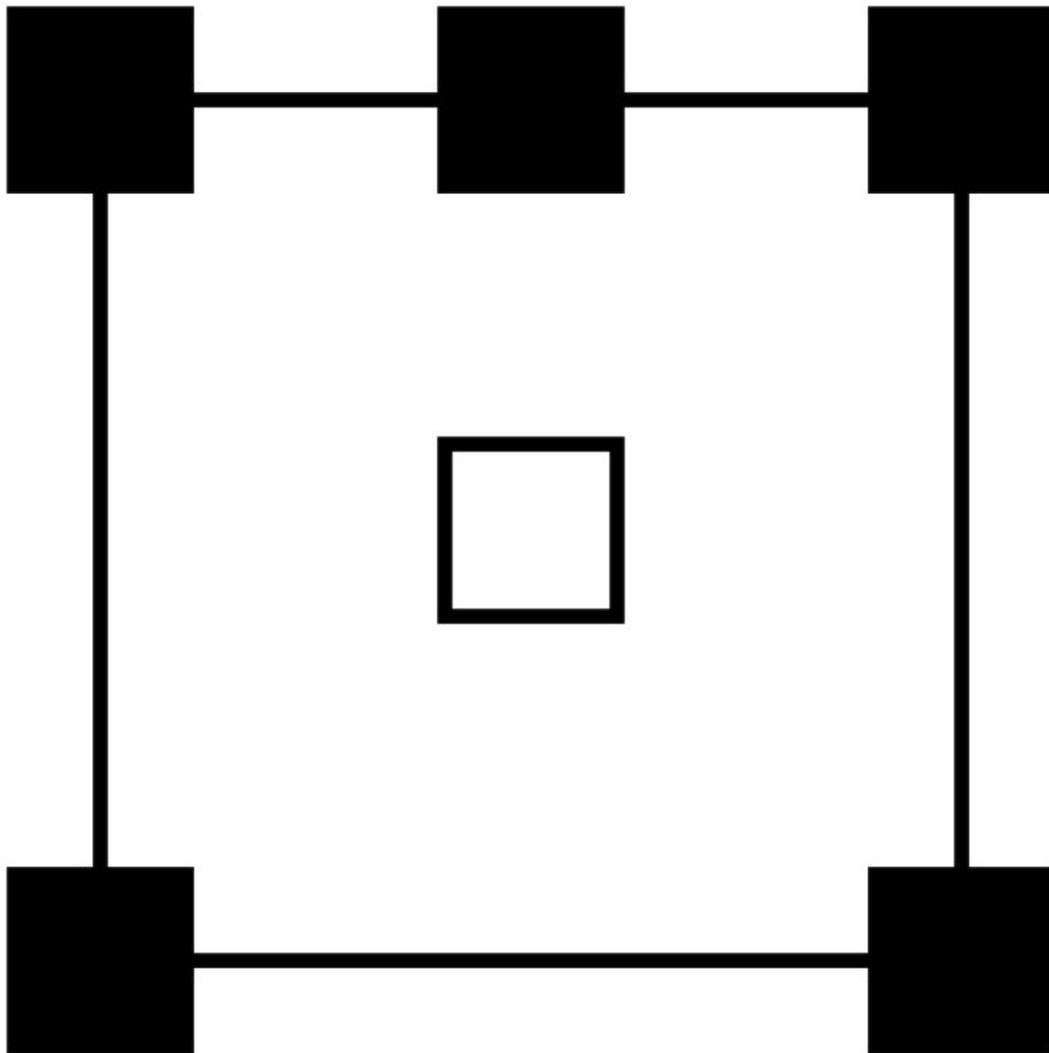
Windows Media Audio Professional (WMA Pro) is an improved lossy codec closely related to WMA standard. It retains most of the same general coding features, but also features improved entropy coding and quantization strategies as well as more efficient stereo coding. Notably, many of the WMA standard's low bitrate features have been removed, as the core codec is designed for efficient coding at most bitrates. Its main competitors include AAC, HE-AAC, Vorbis, Dolby Digital, and DTS. It can support audio resolutions of up to 96 kHz and up to eight discrete channels (7.1 channel surround). WMA Pro also supports dynamic range compression, which reduces the volume difference between the loudest and quietest sounds in the audio track. According to Microsoft's Amir Majidimehr, WMA Pro can technically go beyond 7.1 surround sound and support "an unlimited number of channels."

The codec's bit stream syntax was frozen at the first version, WMA 9 Pro. Later versions of WMA Pro introduced low-bit rate encoding, low-delay audio, frequency interpolation mode, and an expanded range of sampling rate and bit-depth encoding options. A WMA 10 Pro file compressed with frequency interpolation mode comprises a WMA 9 Pro track encoded at half the original sampling rate, which is then restored using a new compression algorithm. In this situation, WMA 9 Pro players which have not been updated to the WMA 10 Pro codec can only decode the lower quality WMA 9 Pro stream. Starting with WMA 10 Pro, eight channel encoding starts at 128 kbit/s, and tracks can be encoded at the native audio CD resolution (44.1 kHz, 16-bit), previously the domain of WMA Standard.

Despite a growing number of supported devices and its superiority over WMA, WMA Pro still has little hardware and software support. Some notable exceptions to this are the Microsoft Zune (limited to stereo), Xbox 360, Windows Mobile-powered devices with Windows Media Player 10 Mobile, newer Toshiba Gigabeat and Motorola devices, and devices running recent versions of the Rockbox alternative firmware. In addition, WMA Pro is a requirement for the WMV HD certification program. On the software side, Verizon utilizes WMA 10 Pro for its V CAST Music Service, and Windows Media Player 11 has promoted the codec as an alternative to WMA for copying audio CD tracks. WMA Pro is supported in Silverlight as of version 2 (though only in stereo mode). In the absence of the appropriate audio hardware, WMA Pro can automatically downmix multichannel audio to stereo or mono, and 24-bit resolution to 16-bit during playback.

A notable example of WMA Pro being used instead of WMA Standard is the NBC Olympics website which uses WMA 10 Pro in its low-bitrate mode at 48 kbit/s.

Windows Media Audio Lossless



Label for 5.1 surround sound, the maximum channel configuration for Windows Media Audio Lossless.

Windows Media Audio Lossless (WMA Lossless) is a lossless audio codec that competes with ATRAC Advanced Lossless, Dolby TrueHD, DTS-HD Master Audio, Apple Lossless, Shorten, Monkey's Audio, FLAC, and WavPack (the last two have the advantage of being open source software and available for nearly any operating system). Designed for archival purposes, it compresses audio signals without loss of quality from the original using VBR. When decompressed, the audio signal is an exact replica of the original. The first version of the codec, WMA 9 Lossless, and its revisions support up to 96 kHz, 24-bit audio for up to 6 discrete channels (5.1 channel surround) with dynamic

range compression control. The typical compression ratio for music varies between 1.7:1 and 3:1.

Hardware support for the codec is available on the Cowon A3 , Bang & Olufsen Serenata, Sony Walkman NWZ-A and NWZ-S series, Zune 4, 8, 80 30, Zune 120 (with firmware version 2.2 or later) and the new Zune HD, Xbox 360, Windows Mobile-powered devices with Windows Media Player 10 Mobile, Toshiba Gigabeat S and V models, Toshiba T-400, the Meizu M3, and Best Buy's Insignia NS-DV, Pilot, and Sport music players. Contrary to some claims, the Archos make of media devices do not support WMA Lossless, nor does the SONOS system . Like WMA Standard, WMA Lossless is being used by a few online stores to distribute music online. Similar to WMA Pro, the WMA Lossless decoder can perform downmixing when capable audio hardware is not present .

Windows Media Audio Voice

Windows Media Audio Voice (WMA Voice) is a lossy audio codec that competes with Speex (used in Microsoft's own Xbox Live online service), ACELP, and other codecs. Designed for low-bandwidth, voice playback applications, it employs low-pass and high-pass filtering of sound outside the human speech frequency range to achieve higher compression efficiency than WMA. It can automatically detect sections of an audio track containing both voice and music and use the standard WMA compression algorithm instead. WMA Voice supports up to 22.05 kHz for a single channel (mono) only. Encoding is limited to constant bit rate (CBR) and up to 20 kbit/s. The first and only version of the codec is WMA 9 Voice.

Windows Mobile-powered devices with Windows Media Player 10 Mobile have native support for WMA 9 Voice playback. In addition, BBC World Service has employed WMA Voice for its Internet radio streaming service.

Sound quality

Microsoft claims that audio encoded with WMA sounds better than MP3 at the same bit rate; Microsoft also claims that audio encoded with WMA at lower bit rates sound better than MP3 at higher bit rates. Double blind listening tests with other lossy audio codecs have shown varying results, from failure to support Microsoft's claims about its superior quality to supremacy over other codecs. One independent test conducted in May 2004 at 128 kbit/s showed that WMA was roughly equivalent to LAME MP3; inferior to AAC and Vorbis; and superior to ATRAC3 (software version).

Some conclusions made by recent studies:

- At 32 kbit/s, WMA Standard was noticeably better than LAME MP3, but not better than other modern codecs in a collective, independent test in July 2004.
- At 48 kbit/s, WMA 10 Pro was ranked second after Nero HE-AAC and better than WMA 9.2 in an independent listening test organized and supported by

Sebastian Mares and Hydrogenaudio Forums in December 2006. This test, however, used CBR for WMA 10 Pro and VBR for the other codecs.

- At 64 kbit/s, WMA Pro outperformed Nero HE-AAC in a commissioned, independent listening test performed by the National Software Testing Labs in 2005. Out of 300 participants, "71% of all listeners indicated that WMA Pro was equal to or better than HE AAC."
- At 80 kbit/s and 96 kbit/s, WMA had lower quality than HE-AAC, AAC-LC, and Vorbis; near-equivalent quality to MP3, and better quality than MPC in individual tests done in 2005.
- At 128 kbit/s, there was a four-way tie between aTuV Vorbis, LAME MP3, WMA 9 Pro and AAC in a large scale test in January 2006, with each codec sounding close to the uncompressed music file for most listeners.
- At 768 kbit/s, WMA 9 Pro delivered full-spectrum response at half the bit rate required for DTS in a comparative test done by EDN in October 2003. The test sample was a 48 kHz, 5.1 channel surround audio track.

Criticism of claimed quality

Microsoft's claims of WMA sound quality have frequently drawn complaints. "Some audiophiles challenge Microsoft's claims regarding WMA's quality," according to a published article from EDN. Another article from MP3 Developments wrote that Microsoft's claim about CD-quality audio at 64 kbit/s with WMA was "very far from the truth." At the early stages of WMA's development, a representative from RealNetworks claimed that WMA was a "clear and futile effort by Microsoft to catch up with RealAudio 8."

Microsoft has sometimes claimed that the sound quality of WMA at 64 kbit/s equals or exceeds that of MP3 at 128 kbit/s (both WMA and MP3 are considered near-transparent at 192 kbit/s by most listeners). In a 1999 study funded by Microsoft, National Software Testing Laboratories (NSTL) found that listeners preferred WMA at 64 kbit/s to MP3 at 128 kbit/s (as encoded by MusicMatch Jukebox). However, a September 2003 public listening test conducted by Roberto Amorim found that listeners preferred 128 kbit/s MP3 to 64 kbit/s WMA audio with greater than 99% confidence. This conclusion applied equally to other codecs at the same bitrate, leading him to conclude that:

“ No codec delivers the marketing plot of same quality as MP3 at half the bitrates. ”

It is important to note that both MP3 and WMA encoders have undergone active development and improvement for many years, so their relative quality may change over time.

A July 2007 public listening test by Sebastian Mares found that 64 kbit/s HE-AAC audio (encoded by Nero Digital) was statistically tied with 64 kbit/s WMA Pro audio, in terms of listener preference.

Players

Apart from Windows Media Player, most of the WMA compression formats can be played using ALLPlayer, VLC media player, MPlayer, RealPlayer, Winamp, Zune Software (with certain limitations—DSP plugin support and DirectSound output is disabled using the default WMA plugin), and many other software media players. The Microsoft Zune media management software supports most WMA codecs, but uses a variation of Windows Media DRM which is used by PlaysForSure.

The FFmpeg project has reverse-engineered and re-implemented the WMA codecs (except WMA Lossless) to allow their use on POSIX-compliant operating systems such as Linux. The rockbox project further extended this codec to be suitable for embedded cores, allowing playback on portable MP3 players and cell phones running open source software. RealNetworks has announced plans to support playback of DRM-unprotected WMA files in RealPlayer for Linux. On the Macintosh platform, Microsoft released a PowerPC version of Windows Media Player for Mac OS X in 2003, but further development of the software has ceased. Microsoft currently endorses the third-party Flip4Mac WMA, a QuickTime component that allows Macintosh users to play WMA files in any player that uses the QuickTime framework. Flip4Mac, however, does not currently support the Windows Media Audio Voice codec.

Encoders

Software that can export audio in WMA format include Windows Media Player, Windows Movie Maker, KM Player, Microsoft Expression Encoder, Sony Sound Forge, GOM Player, RealPlayer, Adobe Premiere Pro, Adobe Audition, and Adobe Soundbooth. Microsoft Office OneNote supports encoding in all WMA codecs, Music Morpher. and Windows Media Encoder supports all available bit rate and resolution options as well.

Digital rights management

The WMA codecs are most often used with the ASF container format, which has an optional DRM facility. Windows Media DRM, which can be used in conjunction with WMA, supports time-limited music subscription services such as those offered by unlimited download services, including MTV's URGE, Napster, Rhapsody, Yahoo! Music Unlimited, and Virgin Digital. Windows Media DRM, a component of PlaysForSure and Windows Media Connect, is supported on many modern portable audio devices and streaming media clients such as Roku, SoundBridge, Xbox 360, and Wii. Players that support the WMA format but not Windows Media DRM list protected titles as unplayable.

Chapter 8

Advanced Audio Coding

Advanced Audio Codings



iTunes DRM-restricted AAC file icon

Filename extension	.m4a, .m4b, .m4p, .m4v, .m4r, .3gp, .mp4, .aac
Internet media type	audio/aac, audio/aacp, audio/3gpp, audio/3gpp2, audio/mp4, audio/MP4A-LATM, audio/mpeg4-generic
Initial release	1997
Type of format	Audio compression format, Lossy compression
Contained by	MPEG-4 Part 14, 3GP and 3G2, ISO base media file format and Audio Data Interchange Format (ADIF)
Standard(s)	ISO/IEC 13818-7, ISO/IEC 14496-3

Advanced Audio Coding (AAC) is a standardized, lossy compression and encoding scheme for digital audio. Designed to be the successor of the MP3 format, AAC generally achieves better sound quality than MP3 at similar bit rates.

AAC has been standardized by ISO and IEC, as part of the MPEG-2 and MPEG-4 specifications. The MPEG-2 standard contains several audio coding methods, including

the MP3 coding scheme. AAC is able to include 48 full-bandwidth (up to 96 kHz) audio channels in one stream plus 16 low frequency effects (LFE, limited to 120 Hz) channels, up to 16 "coupling" or dialog channels, and up to 16 data streams. The quality for stereo is satisfactory to modest requirements at 96 kbit/s in joint stereo mode; however, hi-fi transparency demands data rates of at least 128 kbit/s (VBR). The MPEG-2 audio tests showed that AAC meets the requirements referred to as "transparent" for the ITU at 128 kbit/s for stereo, and 320 kbit/s for 5.1 audio.

AAC is also the default or standard audio format for: Apple's iPhone, iPod, iPad, Nintendo DSi, iTunes, DivX Plus Web Player, Sony's PlayStation 3 and is supported by Sony's PlayStation Portable, latest generation of Sony Walkman, phones from Sony Ericsson, the latest S40 and S60 models from Nokia, Android based phones, Nintendo's Wii (with the Photo Channel 1.1 update installed for Wii consoles purchased before late 2007), and the MPEG-4 video standard.

"*High-Efficiency AAC*" is part of digital radio standards like DAB+ and Digital Radio Mondiale, and mobile television standards DVB-H and ATSC-M/H.

History

AAC was developed with the cooperation and contributions of companies including AT&T Bell Laboratories, Fraunhofer IIS, Dolby, Sony Corporation and Nokia, and was officially declared an international standard by the Moving Picture Experts Group in April 1997. MPEG-2 AAC-LC profile consists of a base format very much like AT&T's Perceptual Audio Coding (PAC) coding format, with the addition of temporal noise shaping (TNS), the Dolby Kaiser Window described below, a nonuniform quantizer, and a reworking of the bitstream format to handle up to 16 stereo, 16 mono, 16 LFE, and 16 commentary channels in one bitstream. The Main profile adds a set of recursive predictors that are calculated on each tap of the filterbank. The SSR uses a 4-band PQMF filterbank, with four shorter filterbanks following, in order to allow for scalable sampling rates.

Standardization

It is specified both as *Part 7 of the MPEG-2 standard*, and *Subpart 4 in Part 3 of the MPEG-4 standard*. As such, it can be referred to as **MPEG-2 Part 7** and **MPEG-4 Part 3 AAC** depending on its implementation, however it is most often referred to as **MPEG-2 AAC**, **MPEG-4 AAC**, or **AAC** for short.

AAC was first specified in the standard MPEG-2 Part 7 (known formally as ISO/IEC 13818-7:1997) in 1997 as a new "part" (distinct from ISO/IEC 13818-3, a.k.a. MPEG-2 BC - backwards compatible) in the MPEG-2 family of international standards. It is known as MPEG-2 NBC (Non-Backward Compatible), because it is not compatible with the MPEG-1 Audio formats (MP3, MP2, MP1). It defined three profiles: Low complexity profile (AAC LC), Main profile (AAC Main) and Scalable sampling rate profile (AAC SSR).

It was updated in MPEG-4 Part 3 (MPEG-4 Audio) (known formally as ISO/IEC 14496-3:1999) in 1999. The MPEG-4 Part 3 standard also defined usage of other audio compression formats (a.k.a. Audio Object Types), such as TwinVQ, CELP, HVXC, Text-To-Speech Interface, Structured Audio and others. A notable addition in this version of the AAC standard is Perceptual Noise Substitution (PNS). MPEG-4 Audio is defined in a way that it remains backwards compatible to MPEG-2 AAC. The MPEG-2 Part 7 profiles - AAC LC profile, AAC Main profile and AAC SSR profile are combined with Perceptual Noise Substitution and defined in the MPEG-4 Audio standard as Audio Object Types (using the same names AAC LC, AAC Main and AAC SSR). MPEG-4 Audio Object Types are combined in four MPEG-4 Audio profiles: Main (which includes most of the MPEG-4 Audio Object Types), Scalable (AAC LC, AAC LTP, CELP, HVXC, TwinVQ, Wavetable Synthesis, TTSI), Speech (CELP, HVXC, TTSI) and Low rate Synthesis (Wavetable Synthesis, TTSI). The reference software for MPEG-4 Part 3 is specified in MPEG-4 Part 4 and the conformance bit-streams are specified in MPEG-4 Part 5.

The MPEG-4 Audio Version 2 (ISO/IEC 14496-3:1999/Amd 1:2000) defined new Audio Object Types - the Low Delay AAC (AAC-LD) object type, Bit-Sliced Arithmetic Coding (BSAC) object type, Parametric audio coding using Harmonic and Individual Line plus Noise and Error Resilient (ER) versions of object types. It also defined four new audio profiles: High Quality Audio Profile, Low Delay Audio Profile, Natural Audio Profile and Mobile Audio Internetworking Profile.

The HE-AAC Profile (AAC LC with SBR) and AAC Profile (AAC LC) were first standardized in ISO/IEC 14496-3:2001/Amd 1:2003. The HE-AAC v2 Profile (AAC LC with SBR and Parametric Stereo) was first specified in ISO/IEC 14496-3:2005/Amd 2:2006. The Parametric Stereo audio object type used in HE-AAC v2 was first defined in ISO/IEC 14496-3:2001/Amd 2:2004.

The current version of the AAC standard is defined in ISO/IEC 14496-3:2009.

AAC+ v2 is also standardized by ETSI (European Telecommunications Standards Institute) as TS 102005.

The MPEG-4 Part 3 standard also contains other ways of compressing sound. These include lossless compression formats, synthetic audio and low bit-rate compression formats generally used for speech.

AAC's improvements over MP3

Advanced Audio Coding is designed to be the successor of the *MPEG-1 Audio Layer 3*, known as MP3 format, which was specified by ISO/IEC in 11172-3 (MPEG-1 Audio) and 13818-3 (MPEG-2 Audio).

Blind tests show that AAC demonstrates greater sound quality and transparency than MP3 for files coded at the same bit rate.

Improvements include:

- More sample frequencies (from 8 to 96 kHz) than MP3 (16 to 48 kHz)
- Up to 48 channels (MP3 supports up to two channels in MPEG-1 mode and up to 5.1 channels in MPEG-2 mode)
- Arbitrary bit-rates and variable frame length. Standardized constant bit rate with bit reservoir.
- Higher efficiency and simpler filterbank (rather than MP3's hybrid coding, AAC uses a pure MDCT)
- Higher coding efficiency for stationary signals (AAC uses a blocksize of 1024 or 960 samples, allowing more efficient coding than MP3's 576 sample blocks)
- Higher coding accuracy for transient signals (AAC uses a blocksize of 128 or 120 samples, allowing more accurate coding than MP3's 192 sample blocks)
- Can use Kaiser-Bessel derived window function to eliminate spectral leakage at the expense of widening the main lobe
- Much better handling of audio frequencies above 16 kHz
- More flexible joint stereo (different methods can be used in different frequency ranges)
- Adds additional modules (tools) to increase compression efficiency: TNS, Backwards Prediction, PNS etc... These modules can be combined to constitute different encoding profiles.

Overall, the AAC format allows developers more flexibility to design codecs than MP3 does, and corrects many of the design choices made in the original MPEG-1 audio specification. This increased flexibility often leads to more concurrent encoding strategies and, as a result, to more efficient compression. However, in terms of whether AAC is better than MP3, the advantages of AAC are not entirely decisive, and the MP3 specification, although antiquated, has proven surprisingly robust in spite of considerable flaws. AAC and HE-AAC are better than MP3 at low bit rates (typically less than 128 kilobits per second). This is especially true at very low bit rates where the superior stereo coding, pure MDCT, and more optimal transform window sizes leave MP3 unable to compete. However, as bit rate increases, the efficiency of an audio format becomes less important relative to the efficiency of the encoder's implementation, and the intrinsic advantage AAC holds over MP3 no longer dominates audio quality.

How AAC works

AAC is a wideband audio coding algorithm that exploits two primary coding strategies to dramatically reduce the amount of data needed to represent high-quality digital audio.

1. Signal components that are perceptually irrelevant are discarded;
2. Redundancies in the coded audio signal are eliminated.

The actual encoding process consists of the following steps:

- The signal is converted from time-domain to frequency-domain using forward modified discrete cosine transform (MDCT). This is done by using filter banks that take an appropriate number of time samples and convert them to frequency samples.
- The frequency domain signal is quantized based on a psychoacoustic model and encoded.
- Internal error correction codes are added;
- The signal is stored or transmitted.
- In order to prevent corrupt samples, a modern implementation of the Luhn mod N algorithm is applied to each frame

The MPEG-4 audio standard does not define a single or small set of highly efficient compression schemes but rather a complex toolbox to perform a wide range of operations from low bitrate speech coding to high-quality audio coding and music synthesis.

- The MPEG-4 audio coding algorithm family spans the range from low bitrate speech encoding (down to 2 kbit/s) to high-quality audio coding (at 64 kbit/s per channel and higher).
- AAC offers sampling frequencies between 8 kHz and 96 kHz and any number of channels between 1 and 48.
- In contrast to MP3's hybrid filter bank, AAC uses the modified discrete cosine transform (MDCT) together with the increased window lengths of 1024 or 960 points.

AAC encoders can switch dynamically between a single MDCT block of length 1024 points or 8 blocks of 128 points (or between 960 points and 120 points, respectively).

- If a signal change or a transient occurs, 8 shorter windows of 128/120 points each are chosen for their better temporal resolution.
- By default, the longer 1024-point/960-point window is otherwise used because the increased frequency resolution allows for a more sophisticated psychoacoustic model, resulting in improved coding efficiency.

Modular encoding

AAC takes a modular approach to encoding. Depending on the complexity of the bitstream to be encoded, the desired performance and the acceptable output, implementers may create profiles to define which of a specific set of tools they want to use for a particular application.

The MPEG-2 Part 7 standard (Advanced Audio Coding) was first published in 1997 and offers three default profiles :

- **Low Complexity (LC)** – the simplest and most widely used and supported;

- **Main Profile (Main)** – like the LC profile, with the addition of backwards prediction;
- **Scalable Sample Rate (SSR)** (MPEG-4 AAC-SSR) – a.k.a. *Sample-Rate Scalable (SRS)*;

The MPEG-4 Part 3 standard (MPEG-4 Audio) defined various new compression tools (a.k.a. Audio Object Types) and their usage in brand new profiles. AAC is not used in some of the MPEG-4 Audio profiles. The MPEG-2 Part 7 AAC LC profile, AAC Main profile and AAC SSR profile are combined with Perceptual Noise Substitution and defined in the MPEG-4 Audio standard as Audio Object Types (under the name AAC LC, AAC Main and AAC SSR). These are combined with other Object Types in MPEG-4 Audio profiles. Here is a list of some audio profiles defined in the MPEG-4 standard:

- **Main Audio Profile** – defined in 1999, uses most of the MPEG-4 Audio Object Types (AAC Main, AAC LC, AAC SSR, AAC LTP, AAC Scalable, TwinVQ, CELP, HVXC, TTSI, Main synthesis)
- **Scalable Audio Profile** – defined in 1999, uses AAC LC, AAC LTP, AAC Scalable, TwinVQ, CELP, HVXC, TTSI
- **Speech Audio Profile** – defined in 1999, uses CELP, HVXC, TTSI
- **Synthetic Audio Profile** – defined in 1999, TTSI, Main synthesis
- **High Quality Audio Profile** – defined in 2000, uses AAC LC, AAC LTP, AAC Scalable, CELP, ER AAC LC, ER AAC LTP, ER AAC Scalable, ER CELP
- **Low Delay Audio Profile** – defined in 2000, uses CELP, HVXC, TTSI, ER AAC LD, ER CELP, ER HVXC
- **Mobile Audio Internetworking Profile** – defined in 2000, uses ER AAC LC, ER AAC Scalable, ER TwinVQ, ER BSAC, ER AAC LD
- **AAC Profile** – defined in 2003, uses AAC LC
- **High Efficiency AAC Profile** – defined in 2003, uses AAC LC, SBR
- **HE-AAC v2 Profile** – defined in 2006, uses AAC LC, SBR, PS

(One of many improvements in MPEG-4 Audio is the Object Type - Long Term Prediction (LTP), which is an improvement of the Main profile using a forward predictor with lower computational complexity.)

Depending on the AAC profile and the MP3 encoder, 96 kbit/s AAC can give nearly the same or better perceptual quality as 128 kbit/s MP3.

AAC error protection toolkit

Applying error protection enables error correction up to a certain extent. Error correcting codes are usually applied equally to the whole payload. However, since different parts of an AAC payload show different sensitivity to transmission errors, this would not be a very efficient approach.

The AAC payload can be subdivided into parts with different error sensitivities.

- Independent error correcting codes can be applied to any of these parts using the Error Protection (EP) tool defined in MPEG-4 Audio standard.
- This toolkit provides the error correcting capability to the most sensitive parts of the payload in order to keep the additional overhead low.
- The toolkit is backwardly compatible with simpler and pre-existing AAC decoders. A great deal of the tool kit's error correction functions are based around spreading information about the audio signal more evenly in the datastream.

Error Resilient (ER) AAC

Error Resilience (ER) techniques can be used to make the coding scheme itself more robust against errors.

For AAC, three custom-tailored methods were developed and defined in MPEG-4 Audio

- **Huffman Codeword Reordering (HCR)** to avoid error propagation within spectral data;
- **Virtual Codebooks (VCB11)** to detect serious errors within spectral data;
- **Reversible Variable Length Code (RVLC)** to reduce error propagation within scale factor data.

AAC Low Delay

The **MPEG-4 Low Delay Audio Coder (AAC-LD)** is designed to combine the advantages of perceptual audio coding with the low delay necessary for two-way communication. It is closely derived from the MPEG-2 Advanced Audio Coding (AAC) format.

Licensing and patents

No licenses or payments are required to be able to stream or distribute content in AAC format. This reason alone makes AAC a much more attractive format to distribute content than MP3, particularly for streaming content (such as Internet radio).

However, a patent license is required for all manufacturers or developers of AAC codecs. It is for this reason FOSS implementations such as FFmpeg and FAAC are distributed in source form only, in order to avoid patent infringement.

Extensions and improvements

Some extensions have been added to the first AAC standard (defined in MPEG-2 Part 7 in 1997):

- **Perceptual Noise Substitution (PNS)**, added in MPEG-4 in 1999. It allows the coding of noise as pseudorandom data;

- **Long Term Predictor (LTP)**, added in MPEG-4 in 1999. It is a forward predictor with lower computational complexity.
- **Error Resilience (ER)**, added in MPEG-4 Audio version 2 in 2000, used for transport over error prone channels;
- **AAC-LD (Low Delay)**, defined in 2000, used for real-time conversation applications;
- **High Efficiency AAC (HE-AAC)**, a.k.a. aacPlus v1 or AAC+, the combination of SBR (Spectral Band Replication) and AAC LC; used for low bitrates; defined in 2003;
- **HE-AAC v2**, a.k.a. aacPlus v2 or eAAC+, the combination of Parametric Stereo (PS) and HE-AAC; used for even lower bitrates; defined in 2004 and 2006;
- **MPEG-4 Scalable To Lossless (SLS)**, defined in 2006, can supplement an AAC stream to provide a lossless decoding option, such as in Fraunhofer IIS's "HD-AAC" product;

Container formats

In addition to the MP4, 3GP and other ISO base media file format-based container formats for storage, AAC audio data may be packaged in a more basic format called Audio Data Interchange Format (ADIF), consisting of a single header followed by the raw AAC audio data blocks. Alternatively, it may be packaged in a streaming format called Audio Data Transport Stream (ADTS), consisting of a series of frames, each frame having a header followed by the AAC audio data. Both formats are defined in MPEG-2 Part 7, but are only considered informative by MPEG-4, so an MPEG-4 decoder does not need to support either format. Two more formats are defined in MPEG-4 Part 3: Low-overhead MPEG-4 Audio Transport Multiplex (LATM), which provides a way to combine separate audio payloads, and Low Overhead Audio Stream (LOAS), a self-synchronizing streaming format.

Products that support AAC

HDTV Standards

Japanese ISDB-T

In December 2003, Japan started broadcasting terrestrial DTV ISDB-T standard that implements MPEG-2 video and MPEG-2 AAC audio. In April 2006 Japan started broadcasting the ISDB-T mobile sub-program, called 1seg, that was the first implementation of video H.264/AVC with audio HE-AAC in Terrestrial HDTV broadcasting service on the planet.

International ISDB-Tb

In December 2007, Brazil started broadcasting terrestrial DTV standard called International ISDB-Tb that implements video coding H.264/AVC with audio AAC-LC

on main program(single or multi) and video H.264/AVC with audio HE-AACv2 in the lseg mobile sub-program.

DVB

The ETSI, the standards governing body for the DVB suite, supports AAC, HE-AAC and HE-AAC v2 audio coding in DVB applications since at least 2004. DVB broadcasts which use the h.264 codec for video normally use the HE-AAC codec for audio.

Hardware

iTunes and iPod

In April 2003, Apple Computer brought mainstream attention to AAC by announcing that its iTunes and iPod products would support songs in MPEG-4 AAC format (via a firmware update for older iPods). Customers could download music in a closed-source Digital Rights Management (DRM)-restricted form of AAC via the iTunes Store or create files without DRM from their own CDs using iTunes. In later years, Apple began offering music videos and movies, which also use AAC for audio encoding.

On May 29, 2007, Apple began selling songs and music videos free of DRM from participating record labels. These files mostly adhere to the AAC standard and are playable on many non-Apple products but they do include custom iTunes information such as album artwork and a purchase receipt, so as to identify the customer in case the file is leaked out onto peer-to-peer networks. It is possible, however, to remove these custom tags to restore interoperability with players that conform strictly to the AAC specification. As of January 6, 2009, nearly all music on the iTunes Store became DRM-free, with the remainder becoming DRM-free by the end of March 2009.

iTunes supports a "Variable bit rate" (VBR) encoding option which encodes AAC tracks in an "Average bit rate" (ABR) scheme. As of September 2009, Apple has added support for HE-AAC (which is fully part of the MP4 standard) but iTunes still lacks support for true VBR encoding. The underlying QuickTime API does offer a true VBR encoding profile however.

Other portable players

- **Archos**
- **Creative Zen Portable**
- **Microsoft Zune**
- **SanDisk Sansa** (some models)
- **Sony PlayStation Portable (PSP)** with firmware 2.0 or greater
- **Sony Walkman**
- **Nintendo DSi**
- **Any portable player that fully supports the Rockbox third party firmware**

Mobile phones

For a number of years, many mobile phones from manufacturers such as Nokia, Motorola, Samsung, Sony Ericsson, BenQ-Siemens and Philips have supported AAC playback. The first such phone was the Nokia 5510 released in 2002 which also plays MP3s. However, this phone was a commercial failure and such phones with integrated music players did not gain mainstream popularity until 2005 when the trend of having AAC as well as MP3 support continued. Most new smartphones and music-themed phones support playback of these formats.

- **Sony Ericsson** phones support various AAC formats in MP4 container. AAC-LC is supported in all phones beginning with K700, phones beginning with W550 have support of HE-AAC. The latest devices such as the P990, K610, W890i and later support HE-AAC v2.
- **Nokia XpressMusic** and other new generation Nokia multimedia phones like N- and E-Series: also support AAC format in LC, HE, M4A and HEv2 profiles
- **BlackBerry**: RIM's latest series of Smartphones such as the 8100 ("Pearl"), 9500 ("Storm") and 8800 support AAC.
- **Apple's iPhone** supports AAC and FairPlay protected AAC files formerly used as the default encoding format in the iTunes store until the removal of DRM restrictions in March 2009.
- The **Motorola Droid** Family supports AAC along with several other audio codecs.
- The HTC Dream (Also known as the T-Mobile G1) is described as supporting certain subset of the full AAC format. As of 2009-04-13 at least several forms of AAC files played while others did not play.

Other devices

- **Apple's iPad**: Supports AAC and FairPlay protected AAC files used as the default encoding format in the iTunes store.
- **Palm OS PDAs**: Many Palm OS based PDAs and smartphones can play AAC and HE-AAC with the 3rd party software Pocket Tunes. Version 4.0, released in December 2006, added support for native AAC and HE-AAC files. The AAC codec for TCPMP, a popular video player, was withdrawn after version 0.66 due to patent issues, but can still be downloaded from sites other than corecodec.org. CorePlayer, the commercial follow-on to TCPMP, includes AAC support. Other PalmOS programs supporting AAC include Kinoma Player and AeroPlayer.
- **Microsoft Windows Mobile** platforms support AAC either by the native Windows Media Player or by third-party products (TCPMP, CorePlayer)
- **Epson** supports AAC playback in the P-2000 and P-4000 Multimedia/Photo Storage Viewers. This support is not available with their older models, however.
- The **Sony Reader** portable eBook plays M4A files containing AAC, and displays metadata created by iTunes. Other Sony products, including the A and E series Network Walkmans, support AAC with firmware updates (released May 2006) while the S series supports it out of the box.

- Nearly every major car stereo manufacturer offers models that will play back .m4a files recorded onto CD in a data format. This includes Pioneer, Sony, Alpine, Kenwood, Clarion, Panasonic, and JVC.
- The **Sonos Digital Media Player** supports playback of AAC files.
- The **Roku SoundBridge** network audio player supports playback of AAC encoded files.
- The **Squeezebox** network audio player (made by Slim Devices, a Logitech company) supports playback of AAC files.
- The *PlayStation 3* supports encoding and decoding of AAC files.
- The *Xbox 360* supports streaming of AAC through the Zune software, and off supported iPods connected through the USB port
- The **Wii** video game console supports AAC files through version 1.1 of the Photo Channel as of December 11, 2007. All AAC profiles and bitrates are supported as long as it is in the .m4a file extension. This update removed MP3 compatibility, but users who have installed this may freely downgrade to the old version if they wish.
- The **Livescribe Pulse and Echo Smartpens** record and store audio in AAC format. The audio files can be replayed using the pen's integrated speaker, attached headphones, or on a computer using the Livescribe Desktop software. The AAC files are stored in the user's "My Documents" folder of the Windows OS and can be distributed and played without specialized hardware or software from Livescribe.

Software

Flash Player

A very common program supporting AAC playback is Flash player, version 9, update 3 and above. Since flash player is also a browser plugin, it can play AAC files through a browser as well.

The Rockbox open source firmware (available for multiple portable players) also offers support for AAC to varying degrees, depending on the model of player and the AAC profile.

Optional iPod support (playback of unprotected AAC files) for the Xbox 360 is available as a free download from Xbox Live.

Other software media players

Almost all current computer media players include built-in decoders for AAC, or can utilize a library to decode it. On Microsoft Windows, DirectShow can be utilized this way with the corresponding filters to enable AAC playback in any DirectShow based player. Mac OS X supports AAC via the QuickTime libraries. Software player applications of particular note include:

- 3ivx MPEG-4 - A suite of DirectShow and QuickTime plugins which support AAC encoding and/or AAC/ HE-AAC decoding in any DirectShow application
- Audio Transcoder - CD Ripper, audio converter, tag editor for Windows, allows to convert from/to AAC.
- CorePlayer also supports LC and HE AAC.
- Easy CD-DA Extractor for Windows, CD Ripper and audio converter, which includes an AAC encoder that supports LC and HE AAC.
- ffdshow is a free open source DirectShow filter for Microsoft Windows operating systems that uses FAAD2 to support AAC decoding.
- foobar2000 is a freeware audio player for Windows that supports LC and HE AAC.
- K-Multimedia Player also supports AAC.
- KSP Sound Player also supports AAC.
- Media Player Classic Home Cinema
- MPlayer or xine are often used as AAC decoders on Linux or Apple Macintosh.
- RealPlayer includes RealNetworks' RealAudio 10 AAC encoder.
- Songbird for Windows, Linux and Mac OS X supports AAC, including the DRM rights management encoding used for purchased music from the iTunes Store, with a plug-in.
- Sony SonicStage also support AAC.
- VLC media player supports playback and encoding of MP4 and AAC files.
- Winamp for Windows, which includes an AAC encoder that supports LC and HE AAC;
- Windows Media Player 12, released with Windows 7, supports playback of AAC files natively.
- Another Real product, Rhapsody supports the RealAudio AAC codec, in addition to offering subscription tracks encoded with AAC.
- XBMC (**X**Box **M**edia **C**enter) supports both AAC (LC and HE) on modified Xbox game-consoles.
- XMMS supports mp4 playback using a plugin provided by the faad2 library.

Some of these players (e.g., foobar2000, Winamp, and VLC) also support the decoding of ADTS (Audio Data Transport Stream) or MP4-contained AAC streamed over HTTP using the SHOUTcast protocol. Plug-ins for Winamp and foobar2000 enable the creation of such streams.

Nero Digital Audio

In May 2006, Nero AG released an AAC encoding tool free of charge, *Nero Digital Audio* (Nero AAC Codec), which is capable of encoding LC-AAC, HE-AAC and HE-AAC v2 streams. The tool is a Command Line Interface tool only, and a separate utility is included to decode to PCM WAV.

Various tools including the foobar2000 audio player, MediaCoder, MeGUI encoding front end and dBpoweramp can provide a GUI for the encoder.

FAAC and FAAD2

FAAC and FAAD2 stand for Freeware Advanced Audio Coder and Decoder 2 respectively. FAAC supports audio object types LC, Main and LTP. FAAD2 supports audio object types LC, Main, LTP, SBR and PS. Although FAAD2 is free software, FAAC is not free software.

FFmpeg

FFmpeg's libavcodec library contains free software codecs for both encoding and decoding AAC (decoding is experimental).

WWT

Chapter 9

DAISY Digital Talking Book

Digital Accessible Information System, or **DAISY**, is a means of creating digital talking books for people who wish to hear—and navigate—written material presented in an audible format; many such listeners have "print disabilities," including blindness, impaired vision, dyslexia or other issues.

Using DAISY, a talking book format is presented with enabled navigation within a sequential and hierarchical structure consisting of (marked-up) text synchronized with audio.

DAISY 2 was a standard based on XHTML and SMIL . DAISY 3 is also based on XML and is standardized as ANSI/NISO Z39.86-2005.

DAISY assists people who, for different reasons, have problems using regular printed media. DAISY books have the benefits of regular audiobooks, but they are superior because DAISY 2.02 provides up to six embedded "navigation levels" for content (i.e. other objects such as images, graphics, MathML etc.) and for displaying synchronized text to speech. DAISY Multimedia can be a talking book, computerised text or a synchronised presentation of text and audio.

As a result, DAISY books allow the blind listener to navigate an encyclopedia; this is impossible using conventional audio recordings because they lack search and navigation features and they require linear listening . While reading a DAISY book, a reader can go to the next or previous page, chapter or sentence. DAISY is for everyone who needs accessible information and for everyone who loves to read.

The DAISY Consortium has been selected by the National Information Standards Organization (NISO) as the official maintenance agency for the DAISY/NISO Standard, officially, the ANSI/NISO Z39.86, Specifications for the Digital Talking Book, known as DAISY 3.

The DAISY Consortium was founded in 1996 and consists of a growing membership of organizations around the world committed to developing equitable access to information for people who have a print disability.

Specification

A DTB is a collection of electronic files arranged to present information to the target population via alternative media, namely, human or synthetic speech, refreshable Braille, or visual display, e.g., large print. The files comprising the Daisy formats are:

- Package File: Drawn from the Open eBook Publication Structure 1.2. It is a XML1.0 file with a set of metadata describing the DTB, a list of files that make up the DTB (the manifest) and a spine that defines the default reading order of the document. Standard: OEBF * Publication Structure 1.2 (the file extension is “.opf”).
- Textual content file: this document contains the text of the document as an XML1.0 according to a specific DTD (dtbook.dtd).
- Audio Files: human or synthetic speech recordings.
- Image files: for visual displays
- Synchronization files: To synchronize the different media files of a DTB during playback, this standard specifies the use of the SMIL2.0.
- Navigation control file: use the NCX (Navigation Control File for XML applications) to view the document’s hierarchical structure, allowing user to move through the book’s larger division or progressively smaller steps (footnotes, for example).
- Bookmark/Highlight file: support to user-set highlights to which text and audio notes can be applied.
- Resource file:
- Distribution Information File: describes how to map each SMIL file to a specific media unit.

Distribution and playback

DAISY books can be heard on standalone DAISY players, computers using DAISY playback software, mobile phones, and MP3 players (with limited navigation). DAISY books can be distributed on a CD/DVD, memory card or through the Internet.

A computerized text DAISY book can be read using refreshable Braille display or screen-reading software, printed as Braille book on paper, converted to a talking book using synthesised voice or a human narration, and also printed on paper as large print book. In addition, it can be read as large print text on computer screen.

Software players

Software-based players include:

- Read:OutLoud 6
- Read:OutLoud Bookshare Edition
- gh PLAYER, a commercial DAISY player

- AMIS - Adaptive Multimedia Information System - an open-source self-voicing player for Windows that works with several screen readers; developed by the DAISY Consortium
- FSReaderDAISY Player Software for PAC Mate and Desktop
- emerson-reader, an open-source and cross-platform Epub and DAISY player
- AnyDaisy Firefox Extension, by Benetech
- Dorina DAISY Reader (DDReader), a Firefox add-on for DAISY 3.0
- ButtercupReader, a web-based application for DAISY 3 books
- CUCAT Olearia, an open-source DAISY reader for Mac OS X
- Daisy Delight (DAISY 2.02, for Mac OS X and Unix-based systems)
- DAISY Book Reader, GTK Daisy Talking Book reader application
- DAISYPlayer (for Windows; only available in Spanish)
- Daisy Player, an OSS/GPL CLI Linux/Ubuntu Daisy Talking Book reader application
- DaisyWorm (for iPhone, iPod touch and iPad; iOS 3.1 or higher)
- Dolphin EasyReader and EasyReader Express, for Windows

Other relevant software includes:

- Daisy Uppsala Archive Project, server-side system for managing DAISY (Digital Talking Books) files
- Online Daisy Delivery Technology, open-source software to deliver DAISY books online

Hardware players

There are also a wide range of hardware products available that can play DAISY content, usually in a portable form factor. Some of these devices are dedicated to playback of books, while others focus on other functionality, such as PDA or mobile Internet access, and offer DAISY playback as either a feature of the unit or as a software add-on.

A short (incomplete) list of products that have built-in support for DAISY playback includes:

- Victor Reader Stream, a hand-held portable DAISY player for the blind, visually handicapped and print impaired, produced by HumanWare
- Victor Reader Wave, also by HumanWare, is a portable CD player that can play DAISY content from CD media
- BookSense, a similar, smaller unit produced by GW Micro; the advanced XT model features built-in flash memory and Bluetooth headset support for playback, as well as an FM radio
- The National Library Service for the Blind and Physically Handicapped (NLS) in the United States has developed a proprietary DAISY player designed for use by its print-disabled patrons. The player will replace the aging cassette-based distribution system.

Access to materials by the disabled

Since DAISY is often used by people with disabilities, many of the existing organizations which produce accessible versions of copyrighted content are moving to the DAISY standard, and slowly moving away from more traditional methods of distribution such as cassette tape.

In the United States, Recording for the Blind and Dyslexic (RFB&D), BookShare and the National Library Service for the Blind and Physically Handicapped (NLS), among others, are offering content to blind and visually impaired individuals. RFB&D also allows access by those with dyslexia or other disabilities which impair the person's ability to read print. The NLS uses a library methodology, on the basis that the books are loaned (as they traditionally have been, on physical cassette), hence they are able to offer content free of charge, just as any public library can. RFB&D and Bookshare both are subscription-based services, however RFB&D has made their membership free of charge to all qualified individuals in the United States

Content from both the NLS and the RFB&D organizations utilizes proprietary encryption extensions to the DAISY standard. The basic structure of the DAISY definition files remains the same, however, the audio itself, and in some cases certain information tags in the DAISY SMIL files, are encrypted and must be decrypted in order to be read/played back. This is done to prevent unauthorized individuals, such as those who do not have a qualifying disability, from accessing the materials. The organization which offers the content provides a decryption key to the user, which can be installed into a DAISY player to allow decryption. As the encryption schemes are not part of the core DAISY standard, only players which specifically implement the necessary algorithms and key management will be able to access these titles. Bookshare does not use such encryption; the data as downloaded from the server is encrypted using only the user's password, which is used to decrypt the data to an open format, and thus content from Bookshare generally is readable on any compliant DAISY player.

Production

Add-ins or extensions to create DAISY files from office software are also available:

- Microsoft and Sonata Software created a Save as DAISY add-in for Microsoft Word to convert Office Open XML text documents to DAISY XML.
- Odt2DAISY is an extension for OpenOffice.org that exports OpenDocument Text to DAISY XML or to Full DAISY (both XML and audio).

Other tools for DAISY production include:

- the DAISY Pipeline, a cross-platform "open source framework for document- and DTB-related pipelined transformations", developed by the DAISY Consortium ,
- the DAISY Pipeline GUI ,
- PipeOnline, a web interface for the DAISY Pipeline ,

- Daisy Producer, an integrated production management system for Digital Talking Books based on the DAISY Pipeline and liblouis
- Z39.86 DTB Validator, "Zedval": "a Java-based conformance validator for ANSI/NISO Z39.86 Digital Talking Books"
- Dolphin Publisher and Dolphin EasyProducer (commercial products)
- Obi, a free authoring tool

WWT

Chapter 10

Audio Codecs

G.719

G.719 is an ITU-T standard audio codec providing high quality, moderate bit rate (32 to 128 kbit/s) wideband (20 Hz - 20 kHz audio bandwidth, 48 kHz audio sample rate) audio coding at low computational load. It was produced through a collaboration between Polycom and Ericsson.

G.719 incorporates elements of Polycom's Siren22 codec (22 kHz) and Ericsson codec technology, as well as Polycom's Siren7 and Siren14 codecs (G.722.1 and G.722.1 Annex C), which have been used in videoconferencing systems for many years. As ITU-T Recommendation G.719, it was approved on June 13, 2008.

G.719 is optimized for both speech and music. It is based on transform coding with adaptive time-resolution, adaptive bit-allocation and low complexity lattice vector quantization. The computational complexity is quite low (18 floating-point MIPS) for an efficient high-quality compressor. The codec operates on 20 ms frames, and the algorithmic delay end-to-end is 40 ms. The encoder input and decoder output are sampled at 48 kHz.

In addition to the nominal bit rates of 32, 48 and 64 kbit/s, the G.719 codec has an inherent feature of flexible rate selection. In fact, it is possible to accommodate any rate between 32 kbit/s and 64 kbit/s by steps of 4 kbit/s. Moreover, the codec can also provide higher rates than 64 kbit/s and up to 128 kbit/s.

Amendment 1 of the ITU-T G.719 specification defined the use of the ISO base media file format (ISO/IEC 14496-12 a.k.a. MPEG-4 Part 12) as container for the G.719 bitstream. It also defined stereo and multichannel use of G.719 bitstreams in the ISO base media file format. It addresses non-conversational use cases of the codec (e.g. call waiting music playback and recording of teleconferencing sessions, voice mail messages). Thus, media file formats such as MP4 (audio/mp4 or video/mp4) and 3GP (audio/3GPP and video/3GPP) can contain G.719-encoded audio.

RFC 5404 defined media type audio/G719.

Licensing

G.719 is licensed by Polycom, Inc. and by Ericsson; both licenses are necessary for use. Polycom licensees also receive the right to use G.722.1 (Siren7, Polycom's 7 kHz codec), and G.722.1 Annex C (Siren14, the 14 kHz equivalent).

G.722

G.722 is a ITU-T standard 7 kHz wideband speech codec operating at 48, 56 and 64 kbit/s. It was approved by ITU-T in November 1988. Technology of the codec is based on sub-band ADPCM (SB-ADPCM).

G.722 sample audio data at a rate of 16 kHz (using 14 bits), double that of traditional telephony interfaces, which results in superior audio quality and clarity.

Other ITU-T 7 kHz wideband codecs include G.722.1 and G.722.2. These codecs are not variants of G.722 and they use different patented compression technologies. G.722.1 is based on Siren codecs and offers lower bit-rate compressions. A more recent G.722.2, also known as AMR-WB ("Adaptive Multirate Wideband") is based on ACELP and offers even lower bit-rate compressions, as well as the ability to quickly adapt to varying compressions as the network topography mutates. In the latter case, bandwidth is automatically conserved when network congestion is high. When congestion returns to a normal level, a lower-compression, higher-quality bitrate is restored.

Applications

G.722 is an ITU standard codec that provides 7 kHz wideband audio at data rates from 48, 56 and 64 kbit/s. This is useful for voice over IP applications, such as on a local area network where network bandwidth is readily available, and offers a significant improvement in speech quality over older narrowband codecs such as G.711, without an excessive increase in implementation complexity. Environments where bandwidth is more constrained may prefer one of the more bit-efficient codecs, such as G.722.1 (Siren7) or G.722.2 (AMR-WB).

G.722 is also widely used by broadcasters for sending commentary grade audio over a single 64kbps ISDN B channel.

RTP Encapsulation

G.722 VoIP is typically carried in RTP payload type 9. Note that IANA records the clock rate for type 9 G.722 as 8 kHz (instead of 16 kHz), RFC3551 clarifies that this is due to a

historical error and is retained in order to maintain backward compatibility. Consequently correct implementations represent the value 8,000 where required but encode and decode audio at 16 kHz.

Whilst G.722 allows for bitrates of 64, 56 and 48 kbit/s, in practice, data is encoded at 64 kbit/s, with bits from the lower sub-band being used to encode auxiliary data. The greater the number of bits allocated to aux data, the lower the bit rate.

G.722.1

G.722.1 is a licensed royalty-free ITU-T standard audio codec providing high quality, moderate bit rate (24 and 32 kbit/s) wideband (50 Hz - 7 kHz audio bandwidth, 16 ksp/s - kilo samples per second) audio coding. It is a partial implementation of Siren 7 codec (which offers bit rates 16, 24, 32 kbit/s) developed by PictureTel Corp. (now Polycom, Inc.). Its official name is *Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss*.

G.722.1 Annex C (or G.722.1C) is a low-complexity extension mode to G.722.1, which doubles the algorithm to permit 14 kHz audio bandwidth using a 32 kHz audio sample rate, at 24, 32, and 48 kbit/s. It is included in the official ITU-T Recommendation G.722.1. The name of this annex is *Annex C - 14 kHz mode at 24, 32, and 48 kbit/s*. It is an implementation of the mono version of Polycom's Siren 14 codec.

G.722.1 is the successor to PT716plus developed by PictureTel Corp. (now Polycom, Inc.), which has been used in videoconferencing systems for many years. As ITU-T Recommendation G.722.1, it was approved on September 30, 1999 after a four-year selection process involving extensive testing. G.722.1/Annex C was approved by ITU-T on May 14, 2005.

G.722.1 is a transform-based compressor that is optimized for both speech and music. The G.722.1 algorithm is based on transform technology, using a Modulated Lapped Transform (MLT). The computational complexity is quite low (5.5 floating-point MIPS) for an efficient high-quality compressor, and the algorithmic delay end-to-end is 40 ms. A 14 kHz (32 ksp/s) extension, G.722.1/Annex C, was approved by ITU-T on May 14, 2005. Also known as the mono version of Siren 14, this extension is also available from Polycom as a royalty-free license.

The numbering of the wideband ITU audio codecs is sometimes confusing. There are three principal codecs, which are unrelated, but all carrying the G.722 label. G.722 is the original 7 kHz codec, using ADPCM and operating at 48 – 64 kbit/s. G.722.1, another 7 kHz codec, operates at half the data rate while delivering comparable or better quality as G.722, but is a transform-based codec. G.722.1 Annex C is very similar to G.722.1, but

provides twice the audio bandwidth, 14kHz. And G.722.2, which operates on wideband speech and delivers very low bitrates, is an ACELP-based algorithm.

G.723

G.723 is a ITU-T standard speech codec using extensions of G.721 providing voice quality covering 300 Hz to 3400 Hz using Adaptive Differential Pulse Code Modulation (ADPCM) to 24 and 40 kbit/s for digital circuit multiplication equipment (DCME) applications. The standard G.723 is **obsolete** and has been superseded by G.726.

Note that this is a completely different codec from G.723.1.

G.723.1

G.723.1 is an audio codec for voice that compresses voice audio in 30 ms frames. An algorithmic look-ahead of 7.5 ms duration means that total algorithmic delay is 37.5 ms. Its official name is *Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s*.

Note that this is a completely different codec from G.723.

There are two bit rates at which G.723.1 can operate:

- 6.3 kbit/s (using 24 byte frames) using a MPC-MLQ algorithm (MOS 3.9)
- 5.3 kbit/s (using 20 byte frames) using an ACELP algorithm (MOS 3.62)

Use

G.723.1 is mostly used in Voice over IP (VoIP) applications due to its low bandwidth requirement. Music or tones such as DTMF or fax tones cannot be transported reliably with this codec, and thus some other method such as G.711 or out-of-band methods should be used to transport these signals. The complexity of the algorithm is below 16 MIPS. 2.2 kilobytes of RAM is needed for codebooks.

G.723.1 is a required audio codec in the H.324 ITU-T recommendation for H.324 terminals offering audio communication. In 3GPP 3G-324M specification support for G.723.1 is not mandatory, but recommended.

Features

- Sampling frequency 8 kHz/16-bit (240 samples for 30 ms frames)
- Fixed bit rate (5.3 kbit/s with 20 byte 30 ms frames, 6.3 kbit/s with 24 byte 30 ms frames)
- Fixed frame size for each rate (5.3 kbit/s with 20 byte 30 ms frames, 6.3 kbit/s with 24 byte 30 ms frames)
- Algorithmic delay is 37.5 ms per frame, with 7,5 ms look-ahead delay
- G.723 is a hybrid speech coder, with high bit rate using Multi-Pulse Maximum Likelihood Quantization (MP-MLQ) and low bit rate using Algebraic Code Excited Linear Prediction (ACELP)
- The complexity of the algorithm is rated at 25, using a relative scale where G.711 is 1 and G.729a is 15.
- G.723.1 Annex A defines 4 byte Silence Insertion Descriptor (SID) frame for Comfort Noise Generation
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.08 for G.723.1 (6.3 kbit/s), compared to 4.45 for G.711 (μ -law)
- PSQM testing under network stress yields Mean Opinion Scores of 3.57 for G.723.1 (6.3 kbit/s), compared to 4.13 for G.711 (μ -law)

G.726

G.726 is an ITU-T ADPCM speech codec standard covering the transmission of voice at rates of 16, 24, 32, and 40 kbit/s. It was introduced to supersede both G.721, which covered ADPCM at 32 kbit/s, and G.723, which described ADPCM for 24 and 40 kbit/s. G.726 also introduced a new 16 kbit/s rate. The four bit rates associated with G.726 are often referred to by the bit size of a sample, which are 2-bits, 3-bits, 4-bits, and 5-bits respectively.

The most commonly used mode is 32 kbit/s, which doubles the usable network capacity by using half the rate of G.711. It is primarily used on international trunks in the phone network. The principal application of 24 and 16 kbit/s channels is for overload channels carrying voice in digital circuit multiplication equipment (DCME). The principal application of 40 kbit/s channels is to carry data modem signals in DCME, especially for modems operating at greater than 4800 kbit/s.

It also is the standard codec used in DECT wireless phone systems and is used on some Canon cameras.

History

G.721 was introduced in 1984, while G.723 was introduced in 1988. They were folded into G.726 in 1990.

G.727 was introduced at the same time as G.726, and includes the same bit rates, but is optimized for packet circuit multiplex equipment (PCME) environment. This is achieved by embedding 2-bit quantizer to 3-bit quantizer and same for the higher modes. This allows dropping of the least significant bit from the bit stream without adverse effects on speech signal.

Features

- Sampling frequency 8 kHz
- 16 kbit/s, 24 kbit/s, 32 kbit/s, 40 kbit/s bit rates available
- Generates a bitstream, therefore frame length is determined by packetization (typically 80 samples for 10 ms frame size)
- Typical algorithmic delay is 0.125 ms, with no look-ahead delay
- G.726 is a waveform speech coder which uses Adaptive Differential Pulse Code Modulation (ADPCM)
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.30 for G.726 (32 kbit/s), compared to 4.45 for G.711 (μ -law)
- PSQM testing under network stress yields Mean Opinion Scores of 3.79 for G.726 (32 kbit/s), compared to 4.13 for G.711 (μ -law)
- 40 kbit/s G.726 can carry 12000 bit/s and slower modem signals, while 32 kbit/s G.726 can carry 2400 bit/s and slower modem signals well and 4800 bit/s with some more degradation than clear channel codecs.

G.728

G.728 is an ITU-T standard for speech coding operating at 16 kbit/s. It is officially described as *Coding of speech at 16 kbit/s using low-delay code excited linear prediction*.

Technology used is LD-CELP, low-delay code excited linear prediction. Delay of the codec is only 5 samples (0.625 ms). The linear prediction is calculated backwards with a 50th order linear predictive coding filter. The excitation is generated with gain scaled VQ. The standard was finished in 1992 in the form of algorithm exact floating point code. In 1994 a bit exact fixed point codec was released. G.728 passes low bit rate modem signals up to 2400 bit/s. Also network signaling goes through. The complexity of the codec is 30 MIPS. 2 kilobytes of RAM is needed for codebooks. Mean opinion score for G.728 is 3.61.

The essence of CELP techniques, which is an analysis-by-synthesis approach to codebook search, is retained in LD-CELP. The LD-CELP however, uses backward adaptation of predictors and gain to achieve an algorithmic delay of 0.625 ms.

RealAudio 28_8 is a reduced-bitrate variant of this standard, using 15.2 kbit/s.

G.729

G.729 is an audio data compression algorithm for voice that compresses digital voice in packets of 10 milliseconds duration. It is officially described as *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)*.

Because of its low bandwidth requirements, G.729 is mostly used in Voice over Internet Protocol (VoIP) applications (such as Skype) where bandwidth must be conserved. Standard G.729 operates at a bit rate of 8 kbit/s, but there are extensions, which provide rates of 6.4 kbit/s (Annex D, F, H, I, C+) and 11.8 kbit/s (Annex E, G, H, I, C+) for worse and better speech quality, respectively.

G.729 has been extended with various features, commonly designated as G.729a and G.729b.

DTMF tones, Fax transmissions, and high-quality audio cannot be transported reliably with this codec. DTMF requires the use of the RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals as specified in RFC 2833.

G.729 Annexes

		G.729 Annexes										
Functionality	-	A	B	C	D	E	F	G	H	I	C+	J
Low complexity		X	X									
Fixed-point	X	X	X		X	X	X	X	X	X		X
Floating-point				X							X	
8 kbit/s	X	X	X	X	X	X	X	X	X	X	X	X
6.4 kbit/s					X		X		X	X	X	
11.8 kbit/s						X		X	X	X	X	
DTX			X				X	X		X	X	
Embedded variable bit rate, wideband												X

G.729 Annex A

G.729a is a compatible extension of G.729, but requires less computational power. This lower complexity, however, bears the cost of marginally reduced speech quality.

G.729a was developed by a consortium of organizations: France Telecom, Mitsubishi Electric Corporation, Nippon Telegraph and Telephone Corporation (NTT), and Université de Sherbrooke.

The features of G.729a are:

- Sampling frequency 8 kHz/16-bit (80 samples for 10 ms frames)
- Fixed bit rate (8 kbit/s 10 ms frames)
- Fixed frame size (10 bytes for 10 ms frame)
- Algorithmic delay is 15 ms per frame, with 5 ms look-ahead delay
- G.729a is a hybrid speech coder which uses Algebraic Code Excited Linear Prediction (ACELP)
- The complexity of the algorithm is rated at 15, using a relative scale where G.711 is 1 and G.723.1 is 25.
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.04 for G.729a, compared to 4.45 for G.711 (u-law)
- PSQM testing under network stress yields Mean Opinion Scores of 3.51 for G.729a, compared to 4.13 for G.711 (u-law)

G.729 Annex B

G.729 has been extended in Annex B (G.729b) which provides a silence compression method that enables a voice activity detection (VAD) module. It is used to detect voice activity in the signal. It also includes a discontinuous transmission (DTX) module which decides on updating the background noise parameters for non speech (noisy frames). It uses 2-byte Silence Insertion Descriptor (SID) frames transmitted to initiate comfort noise generation (CNG). If transmission is stopped, and the link goes quiet because of no speech, the receiving side may assume that the link has been cut. By inserting comfort noise, analog hiss is simulated digitally during silence to assure the receiver that the link is active and operational.

Other extensions

Recently, G.729 has been extended (with Annex J) to provide support for wideband speech and audio coding, i.e., the transmitted acoustic frequency range is extended to 50 Hz - 7 kHz. The respective extension to G.729 is referred to as G.729.1. The G.729.1 coder is hierarchically organized: Its bit rate and the obtained quality are adjustable by simple bitstream truncation.

G.729.1

G.729.1 is an 8-32 kbit/s embedded speech and audio codec providing bitstream interoperability with G.729, G.729 Annex A and G.729 Annex B. Its official name is *G.729-based embedded variable bit rate codec: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729*.

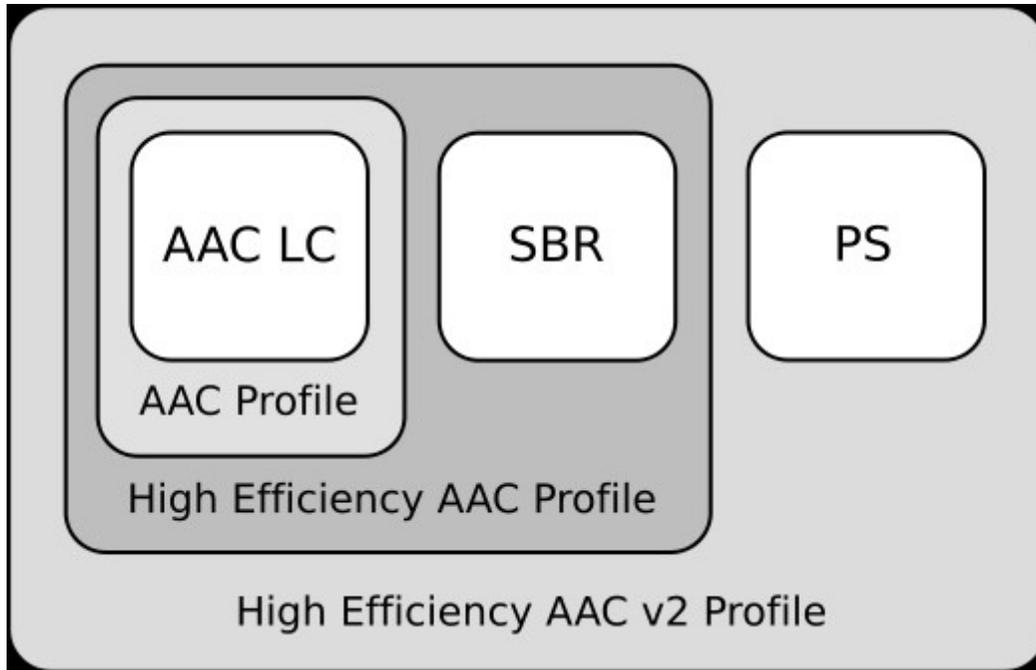
This codec has been designed to provide better quality and more flexibility than the existing ITU-T G.729 speech coding standard. G.729.1 is scalable in bit rate, acoustic bandwidth and complexity. In addition it offers various encoder and decoder modes, including the support of both 8 and 16 kHz input/output sampling frequency, compatibility with G.729B, and reduced algorithmic delay. The bitstream of G.729.1 is structured into 12 hierarchical layers. The first layer (or core layer) at 8 kbit/s follows the G.729 format. The second layer (adds 4 kbit/s for a total of 12 kbit/s) is a narrowband enhancement layer. The third layer (2 kbit/s for a total of 14 kbit/s) is a bandwidth extension layer. Further layers (in 2 kbit/s steps) are wideband enhancement layers. The G.729.1 output bandwidth is 50-4000 Hz at 8 and 12 kbit/s, and 50-7000 Hz from 14 to 32 kbit/s. G.729.1 is also known as G.729 Annex J and G.729EV where EV stands for Embedded Variable (bit rate).

The G.729.1 algorithm is based on a three-stage coding structure: embedded Code-excited linear prediction (CELP) coding of the lower band (50-4000 Hz), parametric coding of the higher band (4000-7000 Hz) by Time-Domain Bandwidth Extension (TDBWE), and enhancement of the full band (50-7000 Hz) by a predictive transform coding technique referred to as Time-Domain Aliasing Cancellation (TDAC).

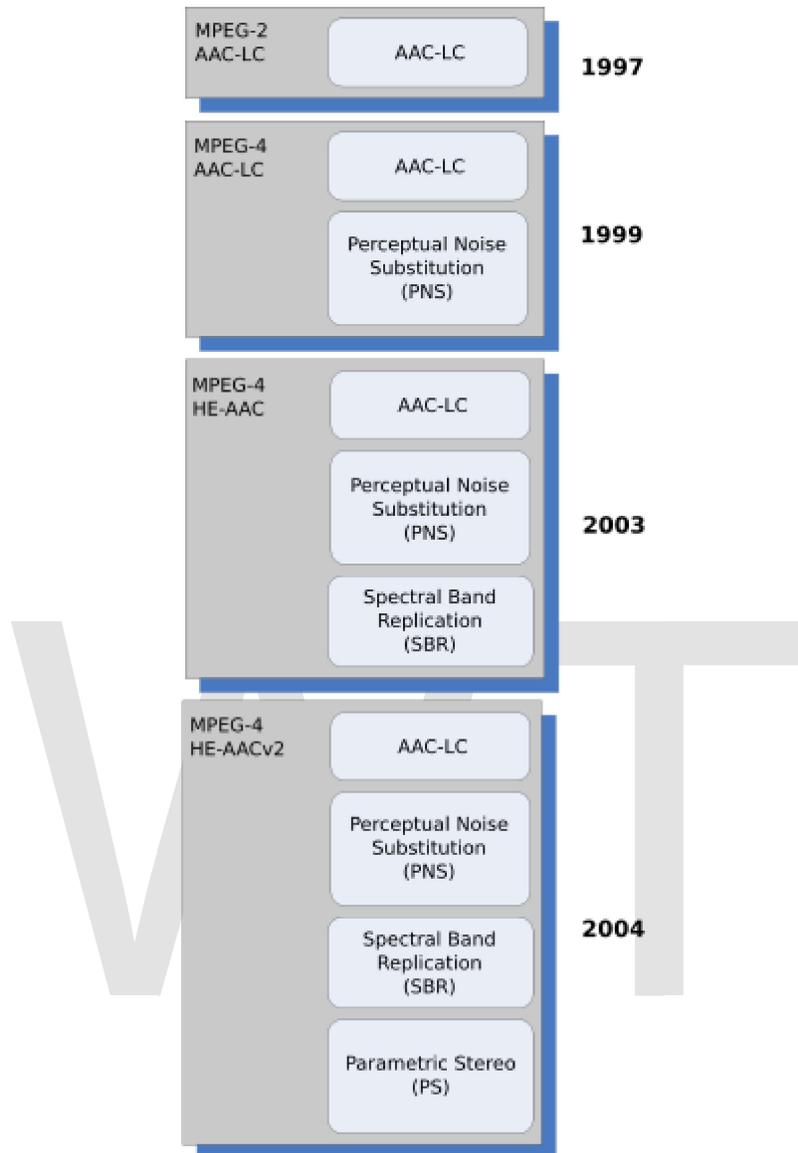
High-Efficiency Advanced Audio Coding

High-Efficiency Advanced Audio Coding

Internet media type	audio/aacp, audio/3gpp, audio/3gpp2
Developed by	ISO
Type of format	Audio compression format
Contained by	3GP, MP4, .dvb, etc.
Extended from	AAC
Standard(s)	ISO/IEC 14496-3



Hierarchical structure of AAC Profile, HE-AAC Profile and HE-AAC v2 Profile, and compatibility between them. The HE-AAC Profile decoder is fully capable of decoding any AAC Profile stream. Similarly the HE-AAC v2 decoder can handle all HE-AAC Profile streams as well as all AAC Profile streams. Based on the MPEG-4 Part 3 technical specification.



Evolution from MPEG-2 AAC LC (Low Complexity) Profile and MPEG-4 AAC LC Object Type to HE-AACv2 Profile.

High-Efficiency Advanced Audio Coding (HE-AAC) is a lossy data compression scheme for digital audio defined as a MPEG-4 Audio profile in ISO/IEC 14496-3. It is an extension of Low Complexity AAC (AAC LC) optimized for low-bitrate applications such as streaming audio. HE-AAC version 1 profile (HE-AAC v1) uses spectral band replication (SBR) to enhance the compression efficiency in the frequency domain. HE-AAC version 2 profile (HE-AAC v2) couples SBR with Parametric Stereo (PS) to enhance the compression efficiency of stereo signals. It is a standardized and improved version of the AACplus codec.

HE-AAC is used in digital radio standards like DAB+ and Digital Radio Mondiale.

History

The HE-AAC version 1 was standardized as a profile of MPEG-4 Audio in 2003 by MPEG and published as part of MPEG-4 in document ISO/IEC 14496-3:2001/Amd 1:2003. The HE-AAC version 2 profile was standardized in 2006 as ISO/IEC 14496-3:2005/Amd 2:2006. A coding tool (Parametric Stereo) used in HE-AAC v2 was standardized in 2004 by MPEG and published in document ISO/IEC 14496-3:2001/Amd 2:2004.

HE-AAC version 2 was also standardized under the name Enhanced aacPlus by 3GPP for 3G UMTS multimedia services in September 2004 (3GPP TS 26.401). It is based on the AAC LC, SBR and Parametric Stereo coding tools defined in the MPEG-4 Audio standard (and it refers to ISO/IEC 14496-3:2001/Amd.1:2003 - Bandwidth Extension, ISO/IEC 14496-3:2001/Amd.2:2004 - Parametric Coding for High Quality Audio, ISO/IEC 14496-3:2001 and ISO/IEC 14496-3:2001/Amd.1:2003/DCOR1). In addition it includes further tools such as error concealment, spline resampler, and stereo-to-mono downmix. It was also published by ETSI as TS 126 401 V6.1.0 in December 2004.

The progenitor of HE-AAC was developed by Coding Technologies under the trade name CT-aacPlus. CT-aacPlus combined MPEG-2 AAC LC with the Coding Technologies invented spectral band replication (SBR). CT-aacPlus is the codec used by XM Radio for their satellite radio service. Subsequently, Coding Technologies submitted their SBR to MPEG as a basis of HE-AAC.

Later, but prior to the standardization of HE AAC v2 by MPEG, Coding Technologies submitted the combination of HE-AAC v1 coupled with Parametric Stereo to 3GPP under the name Enhanced AAC+. As a result, **aacPlus v2** and **eAAC+** are now common trade names that refer to HE-AAC v2. **AAC+** and **aacPlus** are common trade names that refer to HE-AAC v1.

Perceived quality

Scientific testing by the European Broadcasting Union has indicated that HE-AAC at 48 kbit/s was ranked as "Excellent" quality using the MUSHRA scale. MP3 in the same testing received a score less than half that of HE-AAC and was ranked "Poor" using the MUSHRA scale. Data from this testing also indicated that some individuals confused 48 kbit/s encoded material with an uncompressed original.

Testing indicates that material decoded from 64 kbit/s HE-AAC does not yet have similar audio quality to material decoded from MP3 at 128 kbit/s using high quality encoders. The test, taking bitrate distribution and RMSD into account, is a tie between MP3pro, HE AAC and Ogg Vorbis.

Further controlled testing by 3GPP during their revision 6 specification process indicates that HE-AAC and its derivative MPEG-4 HE-AAC v2 provide "Good" audio quality for music at low bit rates (e.g. 24 kbit/s).

MPEG-2 and MPEG-4 AAC LC decoders without SBR support will decode the AAC LC part of the audio, resulting in audio output with only half the sampling frequency, thereby reducing the audio bandwidth.

Support

Encoding

Sony supports HE-AAC encoding since SonicStage version 4. iTunes 9 supports HE-AAC encoding and playback. Nero has released a free-of-charge command line HE-AAC encoder, and also supports HE-AAC inside the Nero software suite. Sorenson Media's Squeeze Compression Suite includes an HE-AACv1 encoder and is available for Mac OS X as well as Windows. The 3GPP consortium released source code of a reference HE-AACv2 encoder that appears to offer competitive quality. Winamp Pro also supports ripping music to HE-AAC. Using a transcoding plugin for Winamp's media library, any file can be transcoded to HE-AAC. XLD, an OS X audio encoding program, offers encoding from any of its supported formats to HE-AAC. Nokia PC Suite may encode audiofiles to eAAC+ format before transmitting them to mobile phone.

Decoding

HE-AAC is supported in the open source FAAD/FAAD2 decoding library (and all players incorporating it): VLC media player, Winamp, foobar2000, Audacious Media Player, and Sony's latest SonicStage version 4. HE-AAC is also used by AOL Radio clients to deliver high-fidelity music at low bitrates.

Adobe's Flash Player 9 supports HE-AACv2.

iTunes 9.2 and iOS4 include full decoding of HE-AAC v2 parametric stereo streams. iTunes 9 thru 9.1, iPhone OS 3.1 and Fall 2009 iPods have support for HE-AAC playback for version 1 with no parametric stereo. Other third-party software featuring support for HE-AAC include FStream, VLC (without metadata/title streaming), Songbird, Snowtape, and recent development versions of Audacious Media Player.

Older versions of Apple iTunes, iPod Touch, and iPhone will play HE-AAC files at reduced fidelity because they ignore the spectral-band replication and parametric stereo information, instead playing them as though they were standard AAC-LC files without the high-frequency, or "treble," information that is only present in the SBR part of the signal. These will report the track length as twice its actual length.

Dolby released Dolby Pulse decoders and encoders in September 2008. HE-AAC v2 is the core of Dolby Pulse so files and streams encoded in Dolby Pulse will playback on

AAC, HE-AAC v1 and v2 decoders. Conversely files and streams encoded in AAC, HE-AAC v1 or v2 will playback on Dolby Pulse decoders.

Dolby Pulse provides the following additional capabilities beyond HE-AAC v2:

- Ability to intelligently generate and insert reversible loudness normalization and dynamic range metadata into the encoded file/stream, this metadata can then be used to optimize the playback experience based on application and/or device.
- Ability to insert custom metadata into the encoded file, and extract this metadata on playback

Dolby has additionally released a PC decoder as an SDK suitable for integration into PC applications requiring Dolby Pulse, HE-AAC or AAC playback capabilities.

Promotion aspects

Commercial trademarks and labeling

HE-AAC is marketed under the trademark aacPlus(tm) by Coding Technologies and under the trademark Nero Digital(tm) by Nero AG. Sony Ericsson, Nokia and Samsung use AAC+ to label support for HE AAC v1 and eAAC+ to label support for HE-AAC v2 on their phones. Motorola uses AAC+ to indicate HE AAC v1 and "AAC+ Enhanced" to indicate HE AAC v2.

Licensing and patents

Companies holding patents for HE AAC have formed a patent pool administered by Via Licensing Corporation to provide a single point of license for product makers.

Patent licenses are required for end-product companies that make hardware or software products that include HE AAC encoders and/or decoders. Unlike the MP3 format, content owners are not required to pay license fees to distribute content in HE AAC.

Standards

HE-AAC profile was first standardized in ISO/IEC 14496-3:2001/Amd 1:2003. HE-AAC v2 profile (HE-AAC with Parametric Stereo) was first specified in ISO/IEC 14496-3:2005/Amd 2:2006. The Parametric Stereo coding tool used by HE-AAC v2 was standardized in 2004 and published as ISO/IEC 14496-3:2001/Amd 2:2004.

The current version of the MPEG-4 Audio (including HE-AAC standards) is published in ISO/IEC 14496-3:2009.

Enhanced aacPlus is required audio compression format in 3GPP technical specifications for 3G UMTS multimedia services and should be supported in IP Multimedia Subsystem

(IMS), Multimedia Messaging Service (MMS), Multimedia Broadcast/Multicast Service (MBMS) and Transparent end-to-end Packet-switched Streaming Service (PSS). HE-AAC version 2 was standardized under the name Enhanced aacPlus by 3GPP for 3G UMTS multimedia services in September 2004 (3GPP TS 26.401).

HE-AAC and HE-AAC v2 audio coding for DVB applications is standardized by TS 101 154. AacPlus v2 by Coding Technologies is also standardized by the ETSI as TS 102 005 for Satellite services to Handheld devices (DVB-SH) below 3 GHz.

In December 2007, Brazil started broadcasting terrestrial DTV standard called International ISDB-Tb that implements video coding H.264 with audio AAC LC on main program (single or multi) and video H.264 with audio HE-AACv2 in the 1Seg mobile sub-program.

Versions

The following is the summary of the different versions of HE-AAC:

Version	Common trade names	Codec feature	Standards
HE-AAC v1	aacPlus v1, eAAC, AAC+, CT-aacPlus	AAC LC + SBR	ISO/IEC 14496-3:2001/Amd 1:2003
HE-AAC v2	aacPlus v2, eAAC+, AAC++, Enhanced AAC+	AAC LC + SBR + PS	ISO/IEC 14496-3:2005/Amd 2:2006

G.711

G.711 is an ITU-T standard for audio companding. It is primarily used in telephony. The standard was released for usage in 1972. Its formal name is *Pulse code modulation (PCM) of voice frequencies*. It is required standard in many technologies, for example in H.320 and H.323 specifications. It can be also used in one of methods for fax communication over IP networks (as defined in T.38 specification).

G.711 represents logarithmic pulse-code modulation (PCM) samples for signals of voice frequencies, sampled at the rate of 8000 samples/second.

G.711.0 (G.711 LLC) - *Lossless compression of G.711 pulse code modulation* was approved by ITU-T in September 2009. It gives as much as 50 percent reduction in bandwidth use.

G.711.1 is an extension to G.711, published as ITU-T Recommendation G.711.1 in March 2008. Its formal name is *Wideband embedded extension for G.711 pulse code modulation*.

G.711, also known as Pulse Code Modulation (PCM), is a very commonly used waveform codec. G.711 uses a sampling rate of 8,000 samples per second, with the tolerance on that rate 50 parts per million (ppm). Non-uniform quantization (logarithmic) with 8 bits is used to represent each sample, resulting in a 64 kbit/s bit rate. There are two slightly different versions; μ -law, which is used primarily in North America, and A-law, which is in use in most other countries outside North America.

Types

G.711 defines two main compression algorithms, the μ -law algorithm (used in North America & Japan) and A-law algorithm (used in Europe and the rest of the world). Both are logarithmic, but A-law was specifically designed to be simpler for a computer to process. The standard also defines a sequence of repeating code values which defines the power level of 0 dB.

The μ -law and A-law algorithms encode 14-bit and 13-bit signed linear PCM samples (respectively) to logarithmic 8-bit samples. Thus, the G.711 encoder will create a 64 kbit/s bitstream for a signal sampled at 8 kHz.

G.711 μ -law tends to give more resolution to higher range signals while G.711 A-law provides more quantization levels at lower signal levels. When using μ -law G.711 in networks where suppression of the all 0 character signal is required, the character signal corresponding to negative input values between decision values numbers 127 and 128 should be 00000010 and the value at the decoder output is -7519. The corresponding decoder output value number is 125.....

G.711 A-Law

A-law encoding thus takes a 13-bit signed linear audio sample as input and converts it to an 8 bit value as follows:

Linear input code Compressed code

s0000000wxyz...	s000wxyz
s0000001wxyz...	s001wxyz
s000001wxyzab...	s010wxyz
s00001wxyzabc...	s011wxyz
s0001wxyzabcd...	s100wxyz
s001wxyzabcde...	s101wxyz
s01wxyzabcdef...	s110wxyz

s1wxyzabcdefg... s111wxyz

Where s is the sign bit, and the ellipsis represents additional low-order bits that are not encoded. So for example, 1000'0000'1010'1111 maps to 1000'1010 (according to the first row of the table), and 0000'0001'1010'1111 maps to 0001'1010 (according to the second).

This can be seen as a floating point number with 4 bits of mantissa and 3 bits of exponent.

In addition, the standard specifies that all resulting even bits are inverted before the octet is transmitted. This is to provide plenty of 0/1 transitions to facilitate the clock recovery process in the PCM receivers. Thus, a silent A-law encoded PCM channel has the 8 bit samples coded 0x55 instead of 0x00 in the octets (or 0xD5 if the sign bit happens to be set).

Note that the ITU define bit 1 to have the value 128 and bit 8 to have the value 1.

The more widely accepted convention has bit 7 = 128 and bit 0 = 1.

Note that when data is sent over E0 (G.703), MSB (signbit) is sent first and LSB is sent last.

G.711 μ -Law

μ -law encoding takes a 14-bit signed linear audio sample as input, increases the magnitude by 32 (binary 100000), and converts it to an 8 bit value as follows:

Linear input code Compressed code

s00000001wxyza... s000wxyz
s0000001wxyzab... s001wxyz
s000001wxyzabc... s010wxyz
s00001wxyzabcd... s011wxyz
s0001wxyzabcde... s100wxyz
s001wxyzabcdef... s101wxyz
s01wxyzabcdefg... s110wxyz
s1wxyzabcdefgh... s111wxyz

Where s is the sign bit, and the ellipsis represents additional low-order bits that are not encoded.

In addition, the standard specifies that all result bits are inverted before the octet is transmitted. Thus, a silent μ -law encoded PCM channel has the 8 bit samples coded 0xFF instead of 0x00 in the octets.

Also the "trick" of adding 32 means μ -law does not encode all 14-bit values; inputs must be within ± 8159 .

Features

- Sampling frequency 8 kHz
- 64 kbit/s bitrate (8 kHz sampling frequency x 8 bits per sample)
- Typical algorithmic delay is 0.125 ms, with no look-ahead delay
- G.711 is a waveform speech coder
- G.711 Appendix I defines a Packet Loss Concealment (PLC) algorithm to help hide transmission losses in a packetized network
- G.711 Appendix II defines a Discontinuous Transmission (DTX) algorithm which uses Voice Activity Detection (VAD) and Comfort Noise Generation (CNG) to reduce bandwidth usage during silence periods
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.45 for G.711 u-law, 4.45 for G.711 a-law
- PSQM testing under network stress yields Mean Opinion Scores of 4.13 for G.711 u-law, 4.11 for G.711 a-law

G.711.1

A recent extension to G.711, G.711.1, allows the addition of narrowband and/or wideband (16000 samples/s) enhancements, each at 25 % of the bitrate of the (included) base G.711 bitstream, leading to data rates of 64, 80 or 96 kbit/s.

G.711.1 is compatible with G.711 at 64 kbit/s, hence an efficient deployment in existing G.711-based voice over IP (VoIP) infrastructures is foreseen. The G.711.1 coder can encode signals at 16 kHz with a bandwidth of 50–7000 Hz at 80 and 96 kbit/s, and for 8-kHz sampling the output may produce signals with a bandwidth ranging from 50 up to 4000 Hz, operating at 64 and 80 kbit/s.

The G.711.1 encoder creates embedded bitstream structured in three layers corresponding to three available bit rates: 64, 80 and 96 kbit/s. The bitstream does not contain any information on which layers are contained, an implementation would require outband signalling on which layers are available. The three G.711.1 layers are: log companded pulse code modulation (PCM) of the lower band including noise feedback, embedded PCM extension with adaptive bit allocation for enhancing the quality of the base layer in the lower band, and weighted vector quantization coding of the higher band based on modified discrete cosine transformation (MDCT).

Two extensions for G.711.1 are planned in 2010: superwideband extension (bandwidth to 14000 Hz) and lossless bitstream compression.